Essays in Environmental and Resource Economics

Thomas O. Michielsen

September 25, 2013

Essays in Environmental and Resource Economics

Proefschrift

ter verkrijging van de graad van doctor aan Tilburg University, op gezag van de rector magnificus, prof. dr. Ph. Eijlander, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit op woensdag 25 september 2013 om 10.15 uur door

Thomas Olivier Michielsen

geboren op 31 januari 1987 te Tiel.

PROMOTIECOMMISSIE:

PROMOTOR:	prof. dr. Reyer Gerlagh
OVERIGE LEDEN:	prof. dr. Aart de Zeeuw
	prof. dr. Sjak Smulders
	prof. dr. Daan van Soest
	prof. dr. Rick van der Ploeg
	prof. dr. Matti Liski

Acknowledgements

I would like to thank a number of people for their help in writing this dissertation. First of all, it has been an enormous pleasure to work with my supervisor, Reyer Gerlagh. Reyer was always willing to interrupt his work to listen to my thoughts, ramblings and conjectures, and respond to them with an astounding intuition and accuracy. Often, after deciphering the essence from my unstructured monologues - no mean feat in itself - he would quickly suggest a result or proof technique that as an unbelieving Thomas I initially discredited, only to come around after a hard day (or two) of work. I also appreciate his informality and wit; working with Reyer was not only learnful, it was also fun. PhD students at other universities have repeatedly told me how lucky I have been to have Reyer as a supervisor, and I could not agree more.

Secondly, I would like to thank Rick van der Ploeg. He has shown incredible kindness and hospitality when inviting me to spend two terms in Oxford, and I enjoyed our conversations about research, academic life and other subjects. I also valued his advice about the job market, as well as his patience with the considerable indecisiveness I have shown in my job search. Thanks for this are also due to Aart de Zeeuw and the members of the job market committee. I mirthfully recall Eline van der Heijden's response when I told her I had agreed to go to Oxford but had not yet signed the contract. Presciently, she warned me that I might change my mind yet again, although my subsequent change of affiliation from the Department of Economics to New College was not as large a switch as to, say, an energy company.

Thijs ten Raa persuaded me to consider a career in academia and start a PhD. His interesting classes in public economics and industrial organization prompted me to seek him out as a thesis supervisor. I am grateful to him for giving me the confidence that I could complete this four-year trajectory, a commodity with which I have not always been naturally endowed.

Arian has been a great office mate during the last three years. Our conversations were great fun, and I learned a lot from your experiences and backgrounds which were different from mine. You were always willing to put up with my quirks and listen to my frustrations and successes and my attempts - with varying degrees of success - to explain my research in everyday language. For making my time as a PhD student more enjoyable I would also like to thank the members of my lunch group, whom I hope did not feel underappreciated because of my frequent seminar attendance.

Lastly, I owe thanks to my friends and family and everyone else who has given me strength and has kept reminding me that the world is a more diverse, nuanced and beautiful place than I sometimes tend to think it is.

Contents

Chapter 1

Introduction

Up until 1800, economic growth has been non-existent or modest at best: between the year 1 and 1820, real income per capita growth was only 0.02% per year on average (Stutz, 2010). The spectacular escape from the Malthusian conjecture of low and stagnant per-capita incomes from the Industrial Revolution onwards has been enabled by a shift away from land-intensive to capital-intensive production techniques, particularly in the energy sector (Hansen and Prescott, 2002). The UK's overreliance on wood as a fuel resulted in serious scarcities in the sixteenth century (Ray, 1979). Though the substitution of coal for wood for domestic heating and industrial use initiated much earlier,[1] the pressure on British wood supplies only abated after the succesful adoption of coke in all stages of the iron smelting process in 1784 (Warde, 2006). Fossil fuels have since offered a cheap, reliable and abundant source of energy. Jevons (1865) extols their significance as follows:

> ”Coal in truth stands not beside but entirely above all other commodities. It is the material energy of the country - the universal aid - the factor in everything we do. With coal almost any feat is possible or easy; without it we are thrown back into the laborious poverty of early times.”

Jevons was one of the first to warn about the exhaustibility of fossil fuels and the implications of a decline in coal production for living standards. The oil crisis in the seventies of the twentieth century drew renewed attention to the prospect of energy shortages. Optimists trust that the market mechanism will appropriately signal the scarcity of natural resources, giving incentives to develop improved extraction techniques, explore new reserves and substitute to alternative energy sources. Pessimists fear that these factors

[1] In the fifteenth century, the use of coal for domestic heating was banned in London to reduce air pollution. By 1615, the paucity of wood had induced a 180 degree turn and it was wood that was banned for this very purpose.

will not be able to compensate for the permanent loss of cheaply accessible supplies.

Over the last few decades, anthropogenic climate change has emerged as one of the biggest global threats. Though the effect of greenhouse gas (GHG) emissions on the global climate, as well as the impacts of higher temperatures and changing weather patterns on human well-being are still uncertain, the consequences of climate change are potentially catastrophic. Fossil fuels are an important contributor to the problem, accounting for 57% of global GHG emissions (IPCC, 2007).

Policymakers across the globe must contend with the two challenges of energy availability and climate change. Scarcity of fossil fuels in the physical sense has not materialized so far: fossil fuel production has not yet peaked, and physically recoverable reserves of coal and unconventional oil and gas are sufficient for the foreseeable future. Economic exhaustion is a more realistic prospect: the deepwater and tar sand oils that are currently being explored are far more costly to extract than reserves in the Middle East. Nonetheless, the recent shale gas boom in the United States and other countries may herald a protracted era of low natural gas prices. But even if these fossil fuels will be more abundantly and cheaply available than renewable alternatives for a long period, continued reliance on them can gravely exacerbate the climate change problem. Climate scientists already advocate a cold turkey style abandonment of fossil fuels (Kharecha and Hansen, 2008).

Limiting climate change and ensuring the continued availability of low-cost energy is complicated by an important market failure. CO2 emissions are a textbook example of a so-called externality: producers and consumers enjoy the full benefit of emitting a unit of carbon, but face only a fraction of the total costs in terms of climate change, because these are shared between all members of current and future generations. This causes individual decision makers to consume more fossil fuel than would be optimal from a social perspective. A carbon price can realign the individual and social objectives: by requiring each emitter to pay a price equal to the burden his emissions impose on contemporaries and future generations, individual decision makers will only emit a unit of carbon if their private benefit exceeds the social cost of emissions.

Regulators that want to impose such a carbon price face limited spatial and temporal jurisdiction however. Free-riding problems make it difficult to form a global climate coalition: each country would like all other countries to participate in an agreement so that it can benefit from their abatement efforts in terms of lower temperatures, but continue under business-as-usual itself (Barrett, 1994). An international treaty offers no panacea: since a supranational authority that can audit emission levels and punish countries that do not abide by their commitments would be at odds with national sovereignty, participating countries still have an incentive to shirk.[2]

The free-rider problem applies even when all agents in the economy are equally concerned about climate change. This dissertation instead focuses on the implications for climate and energy policy when players have different objectives. Fossil fuel producing countries care about export revenues as well as climate change and energy scarcity. Saudi Arabia has been accused of deliberately obstructing climate negotiations (Depledge, 2008), and Canada withdrew from the Kyoto Protocol in 2011. In turn, importers may use environmental policy to capture a share of the fossil fuel rents or for other strategic reasons. European proposals to tax emissions embodied in imports[3] or impose an aviation tax[4] have been criticized by developing countries as 'green protectionism'. The inefficiencies that can result from conflicting objectives and strategic behaviour cannot easily be remedied through multilateral bargaining,[5] because the bargaining outcome may not be enforceable for similar reasons as international environmental agreements.

Aside from different objectives between countries, there are also conflicts between successive regulators or generations. Each regulator may be committed to long-term emission reductions, but prefer the costs of these reductions to incur after her tenure or lifetime. Examples from multilateral declarations abound. The Rio Declaration from 1992 calls upon states to "cooperate in a spirit of global partnership to conserve, protect and restore the health and integrity of the Earth's ecosystem" and "reduce and eliminate unsustainable patterns of production and consumption", yet global carbon emissions from energy increased 48% in the next twenty years.[6] The EU has resolved in 2007 to reduce emissions by 80% in 2050 compared to 1990, but the largest and most costly cuts

[2]See Di Maria et al. (2013) for a non-technical overview of the channels and magnitude of carbon leakage, i.e. the extent to which emissions increase in non-regulated countries following a unilateral emission reduction by a coalition of concerned countries.

[3]India urges rich not to use "green" protectionism, http://www.reuters.com/article/2009/04/07/us-climate-india-bonn-idUSTRE5365FJ20090407, accessed February 11, 2013.

[4]EU's carbon trade plan for aviation is green protectionism, http://news.xinhuanet.com/english/indepth/2011-12/22/c_131321709.htm, accessed February 11, 2013.

[5]In this vein, Harstad (2012) proposes that climate-conscious countries buy the physical coal deposits in less concerned countries, with the intention of leaving them unexploited.

[6]Global carbon emissions rise is far bigger than previous estimates, http://www.guardian.co.uk/environment/2012/jun/21/global-carbon-emissions-record, accessed February 11, 2013.

are reserved for the latter part of this horizon.

The success of energy or climate policy in achieving its aims, whether securing a stable energy supply or reducing cumulative carbon emissions, thus depends on the actions of fossil fuel suppliers and future generations that may not have the same preferences as the policy maker. Effective regulation must take these reactions into account. After highlighting the importance of an abundant energy supply in Chapter 2, this dissertation discusses some implications of conflicting objectives and strategic considerations for energy and climate policy.

Chapter 2 quantifies the effect of energy abundance on industry location. Attracting jobs and profitable industries is a key concern of policy makers, so it is useful to understand what drives firms' decisions to locate in one jurisdiction or another. The chapter specifically asks whether energy-intensive manufacturing sectors have a tendency to locate in US states that have abundant coal, natural gas, oil and hydro reserves. These energy carriers are traded internationally in well-functioning markets but even within the US, local energy prices can differ by a factor three or more. Especially coal and electricity are costly to transport across large distances, and local regulation further contributes to price disparities across states.

Relatively homogeneous jurisdictions such as US states or OECD countries differ more in their energy endowments than in their endowments of capital and skilled labour, which are traditionally considered as the most important factor endowments for firms' location decisions. At the same time, manufacturing sectors are more similar in their capital- and skilled labour requirements than in their energy requirements. The empirical results in Chapter 2 confirm that energy abundance is more important for energy-intensive firms than capital abundance is for capital-intensive firms. States with high coal reserves, such as Montana and Wyoming, mainly attract energy-intensive industries because they have lower electricity prices. Natural gas, oil and hydro endowments also have a direct effect on industry location conditional on energy prices, for example because energy-intensive sectors (e.g. Aluminum and Iron and steel mills) sell part of their output to energy extraction firms, and want to locate close to energy reserves for that reason.

Chapters 3 and 4 deal with incentives of fossil fuel suppliers and their implications for climate and energy policy. In most economic markets, suppliers' decisions how much to produce are more or less separated over time. When firms expect demand to increase in say ten years, they may start investing in additional capacity ahead of time, but today's supply need not adjust. Fossil fuels are exhaustible commodities however: each unit that is extracted today cannot be used to satisfy demand in the future. As a result, their owners constantly compare whether it is more profitable to sell their resources today or in the future. They prefer to sell an additional unit today if the profit from selling an extra unit today (the marginal profit) plus the interest on the proceeds exceeds the marginal profit tomorrow, and vice versa. Along the fossil fuel owners' preferred supply path, marginal profits increase at the interest rate.[7]

Carbon taxes and investments in renewable alternatives typically reduce fossil fuel demand in the medium- and long term more than in the short run. Carbon policies, like the EU Emissions Trading Scheme, often become stricter over time to give producers and consumers some time to adjust their investment decisions, and R&D efforts only result in lower renewable energy prices with a time lag. When climate policies make selling fossil fuels today comparatively more attractive vis a vis selling them in the future, the policies may accelerate fossil fuel extraction. This observation has sparked the fear that anticipated climate policies will worsen rather than solve the climate problem - a 'green paradox' (Sinn, 2008a).[8]

Chapter 3 qualifies this fear by arguing that anticipated carbon taxes are unlikely to increase today's supply of coal and unconventional oil, which are the most potent threat to the global climate. Because these resources are so abundantly available, the tradeoff between extracting a unit today or in the future is less salient than for low-cost oil and natural gas, which are in much more limited supply. Today's supply of coal and unconventional oil primarily depends on today's prices and extraction costs, rather than on expected future market conditions. The chapter derives conditions for which anticipated policies reduce current as well as future emissions. Calibrations suggest that anticipated climate policies are likely to reduce future emissions from coal and unconventional oil much more substantially than they accelerate oil and gas extraction.

Chapter 4 abstracts from climate change and focuses on consumer and producer surplus in the energy market. An importer (for example the OECD) has the ability to

[7]A simple intuition is that a barrel of oil in the ground can be thought of as an asset, which must earn the same return as all other assets in the economy - the interest rate.

[8]Gerlagh and Michielsen (2013) summarize the insights from the economic literature on the green paradox.

develop a substitute for oil by paying an upfront cost, and is both interested in alleviating the physical scarcity of oil and maximizing its share of the oil rent. A monopolistic exporter (say the OPEC) wants to maximize profits while discouraging the OECD from developing the substitute. Because today's oil supply depends on the OPEC's expectations about future conditions, the OECD would like to make a binding promise about the innovation time in order to favourably influence the OPEC's supply schedule. Such a promise would not be credible however, because the OECD has an incentive to renege on its announcement when the oil becomes scarce.

The chapter uses game theoretic methods to derive the OECD's optimal innovation policy and the OPEC's optimal supply rule when neither player can commit to future actions. In equilibrium, the OPEC induces the OECD to delay innovation until the oil is exhausted. By innovating earlier, the OECD loses an important benefit of oil consumption, namely delaying the moment at which the substitute's development costs have to be incurred. Early innovation also causes the OPEC to sell a larger share of its oil reserves just below the substitute price, which is wasteful from the point of view of the OECD. From the OPEC's perspective, the OECD's ability to develop a substitute is equivalent to an already available substitute that is more costly to produce. The chapter raises the possibility that two commonly cited objectives of research in renewable energy - securing a sustainable energy source and obtaining better prices from fossil fuel producers[9] - may actually be at odds with each other.

Chapter 5 explores the implications of conflicting objectives between generations: how can the current generation best adjust its environmental policy when future generations will make a different tradeoff between consumption and preventing an environmental catastrophe, such as severe climate change or large biodiversity loss? Such a disagreement arises naturally when generations share a common concern for preventing a catastrophe, but attach more weight to their own consumption than to the consumption of their descendants. Each generation would then like to implement an austere environmental policy in the future, but enjoy a comparatively high level of consumption today. When today's policy makers realize that their successors will not play their part in such a

[9]US Senator Maria Cantwell for example argued that "[t]ransportation fuel choice could [..] reduce the $200 billion 'monopoly premium' the Department of Energy calculates U.S. consumers currently pay to OPEC and other foreign oil producers each year through excessive petroleum prices." `http://www.cantwell.senate.gov/news/record.cfm?id=334142`, accessed February 12, 2013.

'pollute now, clean up later' plan, the policy that best enacts the current generation's preferences depends on the characteristics of the environmental problem.

Firstly, the chapter considers environmental problems that are caused by a sufficiently scarce pollutant, for example local pollution related to the extraction of an exhaustible resource. In this case, today's generation has an incentive to be conservationist when its descendants have full discretion (in subgame-perfect equilibrium) that is not present when it can fully commit all current and future resource use. Under discretion, future consumption is too rapacious from the current generation's perspective. By reducing its own consumption, the current generation ensures that the resource stock is consumed more smoothly over time, allowing the ecosystem's natural recovery to reduce the probability of a catastrophe.

If the environmental risk is expected to recede in the near future, for example because technological change will make the polluting resource obsolete, today's generation may in contrast have a strategic motive to increase its consumption in subgame-perfect equilibrium. Because the number of future generations that can affect the catastrophe risk is small, the current generation has a direct influence on future decisions. When an increase in today's consumption causes future generations to become more precautionary, today's generation has an incentive to increase its resource use compared to when it has full commitment power, even if this increases the probability of a catastrophe.

Lastly, the chapter analyzes the case in which the polluting resource is abundant and expected to remain essential for a long period. This model has some relevance for climate change if we do not find a substitute for coal and unconventional oil. Here, the catastrophe becomes a self-fulfilling prophecy. Early generations realize that far-future generations will not respect stringent emission ceilings, and that the pollution stock will reach dangerous levels regardless of their actions. Because any mitigation efforts will be undone by future generations, it becomes optimal to continue under business-as-usual. Though each generation has an explicit desire to prevent a catastrophe, generations may act in equilibrium as if they are indifferent about the catastrophe risk. The chapter shows that intergenerational preferences that explicitly value the long-run future can still result in environmental degradation.

Chapter 6 considers a different type of time-inconsistent preferences than Chapter 5, namely hyperbolic discounting: each generation values its own well-being very highly vis a vis their children's well-being, but does not make as sharp a distinction between their children's and their grandchildren's well-being. At the level of an individual, such preferences can explain why people prefer receiving €200 in one year and one month to €150 in one year, but simultaneously favour €150 today over €200 in one month, as well as various kinds of behaviours such as procrastination, addiction, undersaving and lack of exercise. In an intergenerational context, hyperbolic preferences can resolve the tension between high short-term interest rates, which suggest a high degree of impatience, and concerns for far-future generations from stated preference studies and introspection. The positive and normative appeal of these preferences is similar to the preferences in Chapter 5, and future research can shed light which ones are most appropriate for long-lived environmental problems.

The chapter studies the management of a nonrenewable resource with amenity value, for example biodiversity, which provides a range of ecosystem services and is valuable for pharmaceutical research, or the carbon concentration in the atmosphere, which contributes to a hospitable climate. The natural resource provides a stream of benefits when left intact, but can also be irreversibly depleted for immediate economic gain, for example by cutting down the habitat of an endangered species for timber production. Today's generation values the resource's ability to provide amenity value into the far future relatively highly vis a vis the consumption of their immediate descendants. As a result, today's generation may consume the resource if it believes that its descendants will do so otherwise, even if it would prefer to see the resource preserved indefinitely. Today's generation is more inclined to follow a conservationist policy if it is confident that future generations will follow suit.

By contrast, naive policies that ignore the current generation's inability to rule from the grave are less likely to degrade the environment: naive policy makers do not contend with the possibility of future depletion, and thus have no motive to capitalize the resource before their descendants will. The chapter shows that such an unawareness of future preferences need no longer be a blessing when the resource can regenerate naturally. In this case, naive policies can lead to rapacious depletion in the mistaken belief that future generations will restrain themselves to replenish the resource. The chapter provides an example in which a more resilient ecosystem leads to lower welfare if the ecosystem is managed by naive regulators.

Chapter 2

The distribution of energy-intensive sectors in the US[10]

2.1. Introduction

What drives the location of industries? This paper argues that energy is a major determinant. Though coal, natural gas and oil are traded internationally at well-established prices, availability and end-user prices differ substantially across regions. Chemical and metal sectors spend 5-15% of their turnover on energy inputs and benefit greatly from being close to energy reserves, in order to minimize costly transport and to take advantage of energy subsidies. Particularly for coal, transport costs are high compared to the value at the mine,[11] and electricity grids are not designed to handle large volumes of interregional traffic. At the same time, sought-after energy reserves are highly concentrated. The Powder River Basin for example contains more than 40% of coal reserves in the US, the world's most coal-abundant country.

When a select group of industries has a strong incentive to locate near a handful of reserves, variations in energy endowments are an important driver of regional specialization. I test this hypothesis using data on 86 4-digit manufacturing sectors in 50 US states. Coal, natural gas, and to a lesser extent hydro endowments, attract energy-intensive industries. A one standard deviation increase in per capita coal or natural gas endowments increases value added in industries that are more energy-intensive than average by more than 20%. Natural gas, hydro and oil abundance also affect industry location directly

[10] This chapter will also appear as Michielsen (2013a).
[11] Gerking and Hamilton (2008)

once I condition on (instrumented) energy prices, possibly through forward and backward linkages between extractive industries and energy-intensive manufacturing.

The location of manufacturing industries is of great interest to policy makers, and commentators repeatedly highlight the importance of reliable and affordable energy availability for energy-intensive sectors. After the German government's decision to abandon nuclear power in 2011, a leading national newspaper wrote: "What will the new energy age cost us [..] in terms of money and jobs? [..] Energy is the lifeblood of industry, which in turn is the basis of our economy and our prosperity. A stable energy supply is taken for granted [..] and is an enormously important locational advantage when attracting foreign investment."[12] This paper contributes to quantifying the locational advantage that energy abundance provides.

Since the development of the Heckscher-Ohlin model, many studies have tested the factor abundance hypothesis, which states that regions specialize in industries that use their abundant factors.[13] These studies mostly focus on capital and (skilled) labour however.[14] Though capital is a stock input and energy a flow, expenditures on capital and energy as a fraction of turnover (3% and 2%, respectively)[15] suggest that they are an equally important determinant of industry location. Sectors are more similar in their capital- and skilled labour requirements than in their energy requirements, and capital and skilled labour are distributed more evenly across regions than energy reserves. I find that capital abundance is less important for capital-intensive industries than energy abundance is for energy-intensive industries.

To the best of my knowledge, this paper is the first to consider in detail the relation between energy abundance and industry location at the regional level. By focusing on US states instead of countries, I reduce the risk that heterogeneity in technology and consumer tastes dilutes the relation between factor endowments and specialization. Detailed data on energy reserves allow me to insulate effects for four endowments: coal, natural gas, oil and hydro. Furthermore, my primary interest in energy reserves makes

[12] Die Welt, May 30th 2011 (translation by Der Spiegel)

[13] c.f. Bowen et al. (1987), Trefler (1995), Davis et al. (1997) and Romalis (2004)

[14] I list three important exceptions. Hillman and Bullard (1978) assume capital-energy complementarity and conjecture that energy is a source of comparative disadvantage for the US. Ellison and Glaeser (1999) find that US states with low energy prices have a higher activity in energy-intensive sectors, but the direction of causality is unclear. Closely related to the current paper, Gerlagh and Mathys (2011) use data on 14 OECD countries and find that energy-abundant countries produce and export more in energy-intensive industries.

[15] Average during 2001-2009 in US manufacturing, Economic Census

interpretation of the results less prone to endogeneity in the distribution of production factors across states (Schott, 2003). The rest of this paper is organized as follows. Section 2.2 outlines the methodology. Section 2.3 gives an overview of the data. Section 2.4 presents the results, and section 2.5 concludes.

2.2. Methodology

The three main theories of industry location are Ricardo's theory of comparative advantage, the factor abundance hypothesis and the new economic geography literature which emphasizes increasing returns and external economies. While I am foremostly interested in testing whether the factor abundance hypothesis hold for energy carriers, I control for explanations given by the other two theories. This paper is closely related methodologically to Midelfart-Knarvik et al. (2000), Crafts and Mulatu (2005) and Gerlagh and Mathys (2011). In their report on industry location in the EU, Midelfart-Knarvik et al. (2000) interact regional characteristics with sectoral characteristics. If region i has a desirable characteristic j, say an abundance of capital, all sectors will be interested in locating in region i. However, a region's capacity to absorb industries is bounded. The industries that end up in region i are the ones that benefit most from capital abundance and low capital prices, i.e. capital-intensive sectors. Capital-extensive sectors locate somewhere else and are thus underrepresented in region i.

This approach can be applied to production factors, as in the example above, as well as new economic geography (NEG) effects. I include three of these. Industries that have strong forward or backward linkages may agglomerate near large markets, to be close to other producers. I control for this by interacting regional market potential with sectoral forward and backward linkages. Furthermore, industries with large economies of scale may locate in central locations to minimize transport costs. I capture this by interacting market potential with average plant size. I include state-year and sector-year fixed effects.[16] The state-year fixed effects control for any changes in state characteristics

[16] Midelfart-Knarvik et al. (2000) include cutoff levels for state endowments and sectoral intensities in the interaction terms. The interpretation of an endowment cutoff for skilled labour is the endowment level such that an industry's activity does not depend on the skilled labour intensity of the industry. Analogously, the skilled labour intensity cutoff signals the intensity for which industries do not consider the state endowment of skilled labour when making their location decision. Mulatu et al. (2010), studying the effects of environmental regulation on European industry location, present and discuss

Table 2.1: Nomenclature

i	state subscript
s	sector subscript
t	time subscript
j	production factor subscript
l	economic geography subscript
E	set of energy production factors
N	set of non-energy production factors
$VA_{i,s,t}$	value added
$\pi_{j,s,t}$	sectoral factor intensities
$\theta_{j,i,t}$	state factor endowments
$p_{j,i,t}$	state energy prices
$\xi_{j,i}$	state deregulation indicators
$\sigma_{l,s}$	sectoral economic geography characteristics
$\chi_{i,t}$	state market potential

that affect all sectors, which may include changes in the tax code or labour regulation. The sector-year fixed effects absorb any unobserved nationwide sectoral trends, such as price changes for crucial inputs or changes in consumer tastes. I omit sector-state fixed effects as energy reserves and sectoral energy intensities, which enter into the interaction effects of interest, do not vary much over time. To control for this persistence, I cluster error terms by sector-state pair.

I measure industrial activity by value added.[17][18] Table 2.1 describes the notation. I estimate industry location as dependent on a set of production factor interaction terms, three economic geography interaction terms and control variables. The equation I estimate is

$$\ln VA_{i,s,t} = \alpha_{i,t} + \beta_{s,t} + \sum_{j \in \{E \cup N\}} \gamma_j \pi_{j,s,t} \theta_{j,i,t} + \sum_{l} \delta_l \sigma_{l,s} \chi_{i,t} + \epsilon_{i,s,t} \tag{2.1}$$

their estimated cutoff levels in some detail. Because I include fixed effects, I cannot identify these cutoffs.

[17] Midelfart-Knarvik et al. (2000) normalize the left-hand variable by country and sector size. In my specification, size effects are absorbed by the fixed effects.

[18] In a robustness check in the Appendix, I use the log of employment as left-hand variable.

The factor abundance hypothesis predicts that the coefficients on the factor interaction terms γ_j are positive, indicating that energy-intensive industries have a higher value added in energy-abundant states. New economic geography posits that the δ_l coefficients are positive. For ease of interpretation, I discretize the sectoral characteristics such that they are equal to one (zero) if they are higher (lower) than average.[19] As in Mulatu et al. (2010), I normalize state characteristics by their standard deviation. The γ_j coefficients in (2.1) are then comparable across factors. The magnitude of the coefficient on the interaction term for e.g. capital tells us how much variations in capital endowments across states affect the location of industries that are more capital-intensive than average.

I then decompose the effect of energy abundance on the location of energy-intensive industries into a direct effect and an indirect effect through energy prices, as predicted by factor endowment models of trade (c.f. Romalis (2004)). Energy prices are potentially endogenous to industry location, so I instrument for prices using a two-step GMM approach.

$$\ln VA_{i,s,t} = \tilde{\alpha}_{i,t} + \tilde{\beta}_{s,t} + \sum_{j \in \{E \cup N\}} \tilde{\gamma}_j \pi_{j,s,t} \theta_{j,i,t} + \sum_{j \in \{E\}} \tilde{\zeta}_j \pi_{j,s,t} p_{j,i,t} + \sum_{l} \tilde{\delta}_l \sigma_{l,s} \chi_{i,t} + \tilde{\epsilon}_{i,s,t} \tag{2.2}$$

For capital and skilled labour, I preserve the sectoral intensity × state endowment interaction term. For energy, I interact sectoral intensities both with state endowments and state prices. $\tilde{\gamma}_j$ measures the direct effect of energy abundance on energy-intensive sectors, such as forward and backward linkages between the extractive activities and manufacturing sectors (Michaels, 2010), and effects on economic fundamentals such as institutions and infrastructure (c.f. Papyrakis and Gerlagh (2007)). Using the same normalization as for the other interaction terms, the interpretation of $\tilde{\zeta}_j$ is the percentage change in industrial activity in energy-intensive sectors if energy prices increase by one standard deviation.

For electricity and natural gas, I instrument the price interaction terms $\pi_{j,s,t} p_{j,i,t}$ with deregulation interaction terms $\pi_{j,s,t} \xi_{j,i}$. For electricity, I use an indicator whether the electricity sector was deregulated at the start of my sample period.[20] A large part of electricity price differences across states in the eighties and nineties was not related

[19] In the Appendix, I present robustness checks with continuous factor intensities.
[20] Source: EIA (2000).

to fundamentals such as energy abundance and utilization, but to inefficient generation investments and long-term contracts between generators and utilities that stipulated high prices (Joskow, 2000; Borenstein and Bushnell, 2000). Deregulation took place in states in which the gap between regulated prices and the market value of electricity was largest, primarily on the West Coast and in the Northeast. The restructuring proved to be unsuccessful in most states, and did little to curb prices (Blumsack et al., 2006). The deregulation indicator is thus a proxy for long-standing inefficiencies in the electricity sector that, unlike energy reserves, only affect industry location through electricity prices. Because the deregulating states do not have large energy-intensive manufacturing sectors, reverse causality from the composition of the manufacturing industry to deregulation is unlikely.

For natural gas, I employ two similar instruments: indicator variables whether a state adopted a price cap or cost incentive measure for natural gas utilities before the start of my sample period.[21] The utilities, which enjoy natural monopolies, were traditionally subject to rate-of-return regulation. In the nineties and the beginning of the aughts, a number of states introduced price caps and cost-incentive measures. While the two measures are quite different in nature, Hlasny (2011) demonstrates that they were both most likely to be implemented in states with high natural gas prices conditional on geographic characteristics such as natural gas endowments and climate, and in states with high concentration ratios for natural gas distribution. Like in the case of electricity, the reforms did not reduce consumer prices (Hlasny, 2006). Furthermore, I use the population-weighted number of heating degree days[22] as an instrument for natural gas and fuel oil prices. These energy types are widely used as a heating fuel, and temperature-induced variations in demand for heating across states will affect natural gas and fuel oil prices across states.

2.3. Data

I use state-level US panel data containing information on energy reserves, sectoral output and factor inputs, covering a period of 2001-2009. Appendix 2.A.1 contains a full list

[21] Source: Hlasny (2008).
[22] Source: National Oceanic and Atmospheric Administration.

of data sources and definitions. I compute energy abundance as proven reserves per capita. I assume that proven reserves are independent of industrial activity. Though per-capita measures are potentially endogenous, the bias goes against my hypothesis: if energy-abundant states attract more people because they offer better employment opportunities, energy reserves per capita are similar across states. Variations in reserves per capita then have a smaller influence on the location of energy-intensive industries than if reserves per capita are exogenous.

Moreover, area normalizations are distortive as land area is an imperfect indicator of economic potential. Montana and Wyoming, the states with the first and third largest coal reserves, are considerably larger than Illinois, which has the second largest reserves. Comparing these three states using an area-normalized measure of coal abundance, we would predict Illinois to have a comparative advantage in energy-intensive sectors and Montana and Wyoming in energy-extensive, i.e. labour-intensive, sectors. This prediction is implausible as Illinois is more than ten times as populous as the latter two states. Appendix 2.A.2 contains a robustness check in which I define energy abundance as reserves per square mile of land area.

I measure energy intensity as national energy expenditures per employee. The results do not change when measuring energy intensity as energy expenditures per dollar of value added, to ensure that my measure of energy intensity is not clouded by variations in labour intensity. The Energy Information Agency (EIA) has data on four energy endowments: natural gas, coal, oil and hydro.[23] The left-hand variable and factor intensities come from the Annual Survey of Manufactures (ASM) and are observed at the 4-digit NAICS level. For energy, I have data on electricity- and fuel intensity. I use a perpetual inventory approach to construct capital endowments and intensities.

Table 2.2 shows the least and most energy-intensive sectors. Industries that process raw materials, as well as chemical industries, tend to require a lot of energy. Sectors that are further down the production chain, especially those that cater to consumers, are typically energy extensive. Figure 2.1 depicts the sources of electricity generation in in the US. Coal is the predominant source of electricity, accounting for nearly 50%

[23]Hydropower generation is constrained by geographic characteristics, so I regard hydropower capacity as an adequate proxy for the endowment of suitable hydropower generation locations. The location of nuclear power plants may be endogenous, so I do not include this electricity source in the econometric analysis.

Table 2.2: Four-digit NAICS sectors with lowest and highest shares of turnover spent on energy

	Least energy intensive	$\frac{\text{Energy expenditures}}{\text{Turnover}}$		Most energy intensive	$\frac{\text{Energy expenditures}}{\text{Turnover}}$
3341	Computers & peripheral equipment	0.30%	3315	Foundries	4.77%
3122	Tobacco	0.30%	3252	Resin, syn rubber, & artificial syn fibers	4.83%
3361	Motor vehicles	0.39%	3279	Other nonmetallic mineral products	4.85%
3342	Communications equipment	0.45%	3313	Aluminum	7.10%
3343	Audio and video equipment	0.46%	3271	Clay products & refractories	7.26%
3379	Other furniture	0.48%	3251	Basic chemicals	7.31%
3369	Other transportation equipment	0.51%	3311	Iron & steel mills	7.42%
3345	Electronic instruments	0.59%	3272	Glass & glass products	7.62%
3391	Medical equipment & supplies	0.60%	3221	Pulp, paper, & paperboard mills	9.03%
3362	Motor vehicle bodies & trailers	0.60%	3274	Lime and gypsum	13.91%

Shares are averaged over 2001-2009, Annual Survey of Manufactures

Table 2.3: Sample correlation coefficients for energy interaction terms

	(1)	(2)	(3)	(4)	(5)
(1) electricity intensity × coal abundance	1				
(2) electricity intensity × natural gas abundance	0.33	1			
(3) electricity intensity × hydro abundance	0.24	0.11	1		
(4) fuel intensity × natural gas abundance	0.26	0.81	0.07	1	
(5) fuel intensity × oil abundance	0.40	0.62	0.14	0.75	1

of total generation. As natural gas is also an important input, I interact natural gas endowments with electricity intensity as well as with fuel intensity.

Table 2.3 presents the correlation between the energy interaction terms. Multicollinearity between the natural gas and oil interaction terms makes it difficult to disentangle the effect of natural gas and oil endowments on the location of energy-intensive industries. In the Appendix, I use a more disaggregated measure of energy intensity, decomposing fuel intensity into natural gas-, distillate fuel oil- and residual fuel oil intensity. The extra information on energy intensities comes at the cost of observing 3-digit rather than 4-digit sectors.

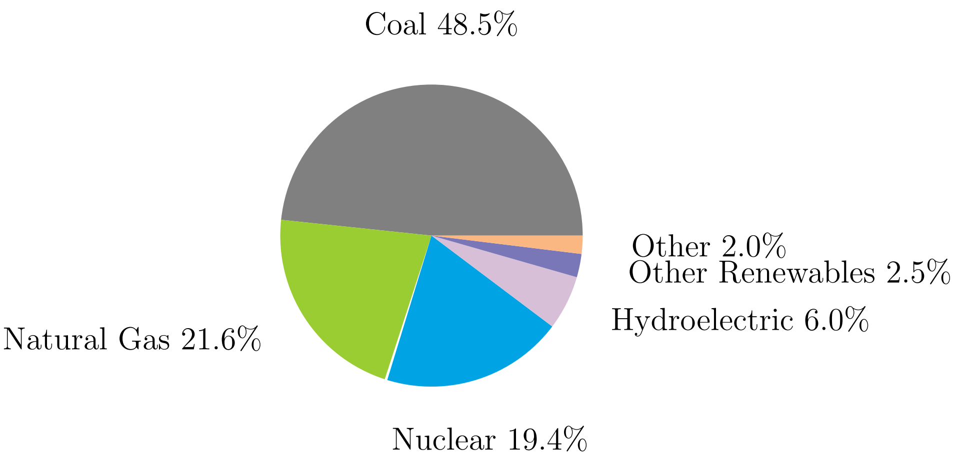

Figure 2.1: Electricity generation in the US by source in 2007 (EIA)

Lastly, Table 2.4 shows the correlation between energy reserves and prices. In accordance with the factor abundance hypothesis, industrial energy prices tend to be lower in states with large energy reserves. Consistent with Figure 2.1, coal and hydro abundance are negatively correlated with electricity prices. The relation between natural gas abun-

Table 2.4: Correlation between state energy abundance (rows) and prices (columns) in 2007

	Electricity	Natural gas	Distillate fuel oil	Residual fuel oil
Coal	−0.26∗	−0.37 ∗ ∗∗	0.18	−0.33 ∗ ∗
Hydro	−0.25∗	−0.04	0.04	−0.20
Natural gas	−0.05	−0.53 ∗ ∗∗	0.00	−0.37 ∗ ∗∗
Oil	−0.03	−0.53 ∗ ∗∗	0.06	−0.39 ∗ ∗∗

[a] Prices are for the industrial sector. Abundance and prices are truncated at the 95th percentile. Asterisks denote significance at the 10% (*), 5% (**) and 1% (***) level. Source: State Energy Data System, EIA.

dance and electricity prices is weak however. Natural gas and oil abundance are strongly negatively correlated with natural gas and residual fuel oil prices respectively. There is no significant relation between distillate fuel oil prices and oil reserves, suggesting that the distillate fuel oil market is more nationally integrated than the residual fuel oil one.

2.4. Results

Table 2.5 presents the results of regression (2.1). The effect of energy endowments on industry location is both statistically and economically significant. A one standard deviation increase in per capita coal or natural gas endowments increases the value added of electricity-intensive sectors by 23%. The effect for hydro is slightly lower, as hydropower constitutes a smaller share of total electricity generation. Natural gas endowments play an even stronger role in the location of fuel-intensive industries: a one standard deviation increase in endowments brings about a 39% increase in value added in fuel-intensive sectors. I find no evidence that oil endowments matter for the location of fuel-intensive industries. The senstivity analyses with finer energy-intensity disaggregation in the Appendix suggest that the negative sign on the oil interaction term is caused by multicollinearity with the natural gas interaction terms. The US possess 4.5% of world conventional natural gas reserves, but only 1.5% of oil reserves.[24] For oil-intensive industries, access to imports may be more important than for natural gas-intensive industries.

[24] International Energy Statistics 2009, EIA

Table 2.5: Location of 4-digit NAICS sectors

	(1)	(2)	(3)
electricity intensity × coal abundance	0.23		0.25
	(0.058)***		(0.061)***
electricity intensity × natural gas abundance	0.25		−0.04
	(0.069)***		(0.074)
electricity intensity × hydro abundance	0.17		0.18
	(0.047)***		(0.046)***
fuel intensity × natural gas abundance		0.39	0.48
		(0.076)***	(0.088)***
fuel intensity × oil abundance		−0.05	−0.16
		(0.104)	(0.112)
capital intensity × capital abundance	0.07	0.07	0.07
	(0.045)	(0.045)	(0.045)
skill intensity × skill abundance	0.05	0.05	0.06
	(0.030)*	(0.030)*	(0.030)*
forward linkages × market potential	0.05	0.01	0.05
	(0.036)	(0.035)	(0.036)
backward linkages × market potential	−0.12	−0.13	−0.12
	(0.035)***	(0.035)***	(0.035)***
scale economies × market potential	−0.07	−0.08	−0.08
	(0.039)*	(0.039)**	(0.039)**
Number of observations	18423	18908	18423
Number of clusters	2971	3071	2971
Adjusted R^2	0.57	0.57	0.57

This table reports coefficient estimates for regression equation (2.1). Variable definitions are given in Tables 2.7 and 2.8. Two-way fixed effects (state-year and sector-year) are included. Error terms are clustered by sector-state pair. Standard errors in parentheses. Asterisks denote significance at the 10% (*), 5% (**) and 1% (***) level. Sector characteristics are equal to one if they are larger than average, and zero if they are smaller than average. State characteristics are divided by their standard deviation.

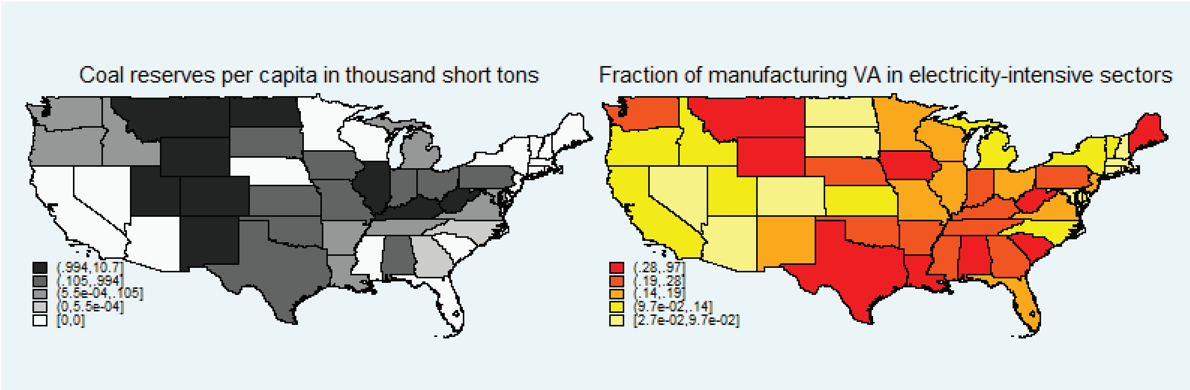

Figure 2.2: Coal endowments per capita (left) and value added in electricity-intensive industries per capita (right), 2007

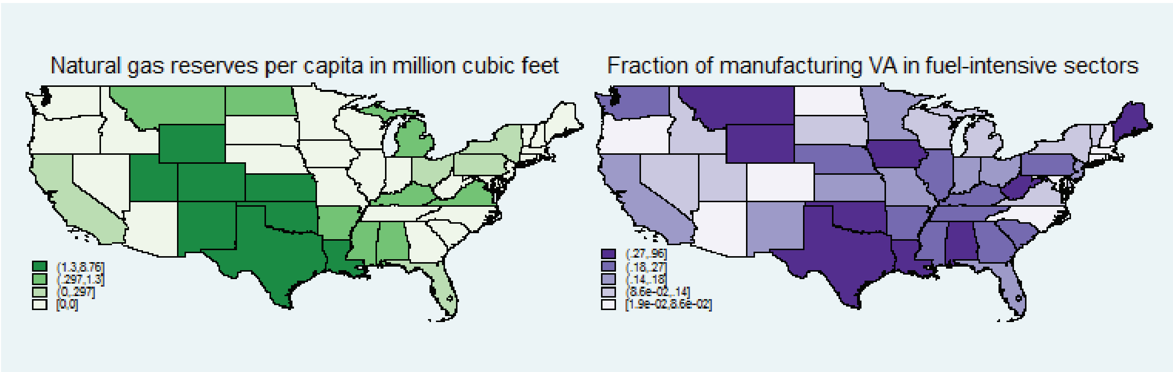

Figure 2.3: Natural gas endowments per capita (left) and value added in fuel-intensive industries per capita (right), 2007

Figure 2.2 illustrates that electricity-intensive industries tend to locate in states with large coal endowments. The left panel shows coal endowments per capita; the right panel value added in electricity-intensive sectors per capita. There are two main coal producing regions in the US: the Appalachians (with Illinois, Kentucky and West Virginia as the most abundant states) and the Western Coal Region (with large reserves in Wyoming, North Dakota and Montana). With the exception of North Dakota, states in these regions also have a high value added in electricity-intensive sectors. Figure 2.3 shows that fuel-intensive sectors are overrepresented in natural-gas abundant Great Plains (Wyoming, Oklahoma and Colorado) and Gulf Coast (Texas and Louisiana) states.

By comparison, the distribution of capital and skilled labour across the US has a much smaller influence on the location of industries that rely on these factors more than average. The coefficients (0.07 and 0.05 respectively) are smaller in absolute value than those of the energy interaction terms, and not significant at the 5% level. Where does

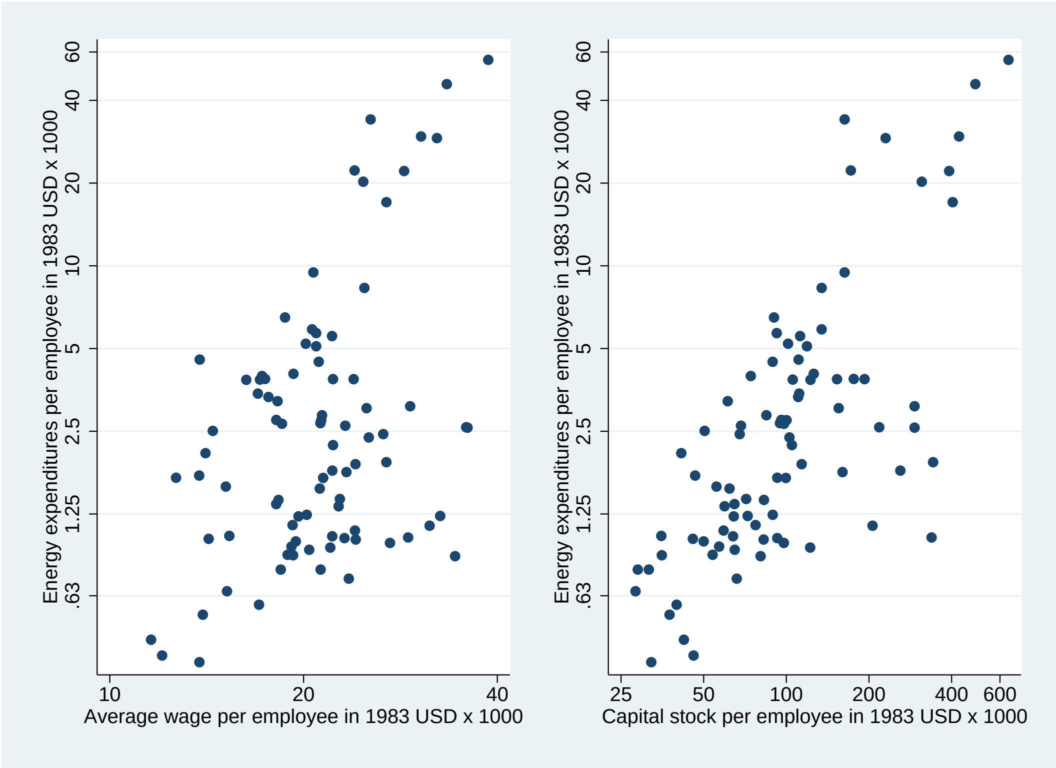

Figure 2.4: Labour, capital and energy intensities in 4-digit US manufacturing sectors. (Source: Economic Census, 2007)

this difference come from? Firstly, the variation in energy intensities across sectors is much higher than the variation in capital- and skill intensities (Figure 2.4). Energy expenditures in the most energy-intensive sector are 152 times as large as in the least energy-intensive sector. The largest difference factors for capital- and skill intensities are 22.8 and 3.3, respectively. For industries that are extremely energy-intensive, being close to energy reserves can be a crucial consideration in the location decision. Because capital- and skill-intensities are much less skewed, locating in capital- or skill-abundant states is not as overriding a concern for capital- and skill-intensive industries.

Secondly, state energy endowments are much more concentrated than capital and skilled labour endowments. Figure 2.5 illustrates this point for natural gas reserves. More than half of states has no natural gas at all, whereas the most abundant state (Wyoming) has 155 times as many reserves per capita as New York. Sectors that require a lot of energy thus have limited options if they want to locate close to energy reserves. Skilled labour and physical capital on the other hand are available in every state. The fraction of adults with a Bachelor's degree or higher differs by a factor 2.2 at most; the largest difference in physical capital per capita is a factor 8.1. As capital and skill

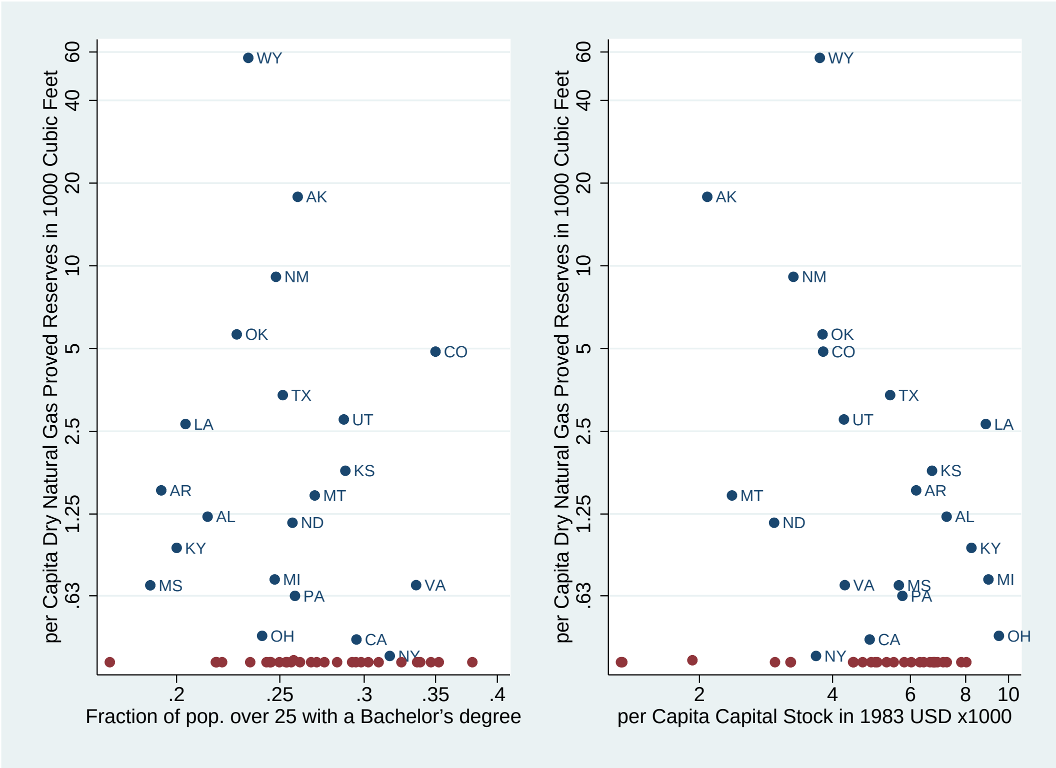

Figure 2.5: Natural gas, capital and skilled labour endowments of US states (Source: EIA and Economic Census, 2007). Dots without labels represent states with little or no gas reserves.

endowments vary much less across states than energy reserves, we can expect variations in endowments to play a smaller role in the location decision for capital- and skill-intensive industries than for energy-intensive industries.

I find no support for the new economic geography hypotheses. Market potential does not significantly affect the location of sectors that have more forward linkages than average. For backward linkages and economies of scale, my findings contradict the NEG prediction: sectors with strong backward linkages and scale economies are under- rather than overrepresented in large markets.

Table 2.6 shows the results when I include both endowment and price interaction terms (for brevity, I do not report the coefficients on the capital, skill and NGE interaction terms). The first-stage IV results are in Table 2.9 in the Appendix. Energy abundance is significant in the first-stage IV regressions. A one standard deviation increase in natural gas abundance is associated with a 38% decrease in natural gas prices. Coal abundance has a stronger influence on electricity prices than hydro or natural gas abundance, which is consistent with the high share of coal-fired electricity generation in Figure 2.1. States

Table 2.6: OLS and GMM-IV coefficient estimates for location of 4-digit NAICS industries with energy intensity and price interaction terms

	(1)		(2)		(3)		(4)	
	OLS	GMM-IV	OLS	GMM-IV	OLS	GMM-IV	OLS	GMM-IV
electricity intensity × electricity price	-0.24	-0.33						
	(0.042)***	(0.094)***						
fuel intensity × natural gas price			-0.18	-0.58				
			(0.045)***	(0.229)**				
fuel intensity × distillate fuel oil price					-0.07	-0.18		
					(0.035)**	(0.088)**		
fuel intensity × residual fuel oil price							-0.07	1.63
							(0.033)**	(225.866)
electricity intensity × coal abundance	0.14	0.08						
	(0.060)**	(0.065)						
electricity intensity × hydro abundance	0.13	0.12						
	(0.047)***	(0.049)**						
electricity intensity × natural gas abundance	0.22	0.19						
	(0.069)***	(0.063)***						
fuel intensity × natural gas abundance			0.28	-0.01				
			(0.062)***	(0.161)				
fuel intensity × oil abundance					0.28	0.24	0.27	0.80
					(0.086)***	(0.071)***	(0.089)***	(67.074)
Number of observations	18423	18423	18908	18908	18908	18908	18908	18908
Number of clusters	2971	2971	3071	3071	3071	3071	3071	3071
Adjusted R^2	0.57	0.56	0.57	-0.01	0.56	0.44	0.56	-0.10
Hansen overidentification p-value				1.00				
Kleibergen-Paap statistic		237.48		19.67		237.79		21.90

This table reports coefficient estimates for regression equation (2.2). All regressions include capital, skill and NGE interaction terms. In the GMM-IV regressions, the energy price interaction terms are instrumented. The excluded instruments are the following. In (1): electricity intensity × electricity deregulation. In (2): fuel intensity× natural gas price cap, fuel intensity× natural gas cost incentive and fuel intensity× heating degree days. In (3) and (4): fuel intensity× heating degree days. Results for the first-stage IV regressions are reported in Table 2.9. Two-way fixed effects (state-year and sector-year) are included. Error terms are clustered by sector-state pair. Standard errors in parentheses. Asterisks denote significance at the 10% (*), 5% (**) and 1% (***) level. Sector characteristics are equal to one if they are larger than average, and zero if they are smaller than average. State characteristics are divided by their standard deviation.

that adopted deregulation are characterized by higher electricity and natural gas prices. The coefficients on the endowment interaction terms in the second stage are smaller than in specification (2.1) without price interaction terms. The results indicate that part of the effect of energy endowments on the location of energy-intensive industries goes through energy prices. Energy abundance causes lower energy prices, which in turn attract energy-intensive manufacturing sectors.

Natural gas, hydro and oil endowments also have a direct effect on the location of energy-intensive industries; for coal, the direct effect becomes insignificant when I instrument for electricity prices. Compared to coal, the former three energy types are characterized by a more capital-intensive extraction or generation process, which may result in stronger linkages with manufacturing.

When I instrument for energy prices, the coefficients on the price interaction terms are larger in absolute value than in the OLS regressions. In the OLS results, a one standard deviation increase in electricity or natural gas prices results in a 24% or 18% decrease in value added in electricity- or fuel-intensive industries, respectively. In the IV results, the decrease is 33% and 58%. This result suggests that the presence of large energy-intensive sectors drives up energy prices in energy-abundant states, ameliorating the effect of energy endowments on prices. This demand effect is stronger than possible negative influences of energy-intensive industries on energy prices, for example through lobbying.

The larger magnitudes of the IV coefficients for price interaction terms may also be caused by measurement error in the industrial energy prices.[25] If energy abundance is better correlated with the true prices than the observed prices are, the IV estimates will be larger in absolute value than the OLS estimates. The large standard deviations in IV regression 4 suggest that the number of heating degree days is not an informative instrument for residual fuel prices.

Appendix 2.A.2 presents the robustness checks. The effect of hydro abundance on specialization is less robust than that of coal and natural gas, but the main conclusions remain unaltered.

[25] Gerlagh and Mathys (2011) propose a method to recover energy prices that is based on the optimality conditions for a Cobb-Douglas production function. They regress observed energy use on a sector- and country fixed effect. The coefficients on the country fixed effects can then be interpreted as a proxy for marginal costs. This approach requires energy use data with variation across both sectors and states, to which I do not have access.

2.5. Conclusion

Though energy is often overlooked in industry location analyses because it is a tradable commodity, it plays a significant role in the distribution of manufacturing sectors. My findings suggest that energy is more important than capital and skilled labour for the location of manufacturing industries in the US. Energy abundance affects industry location indirectly through lower energy prices, but also directly when I condition on prices. The analysis in this paper focuses on reserves as an exogenous source of energy abundance, but policy makers can also influence energy availability, for example through investments in nuclear, solar and wind energy. Considering the strong influence of coal and hydro endowments on the location of electricity-intensive industries, such investments can play an important role in attracting value added and employment in energy-intensive sectors.

2.A. Appendix

2.A.1. Data definitions and sources

Tables 2.7 and 2.8 list the definitions and sources of the state and sector characteristics, respectively. I measure state capital abundance by the manufacturing capital stock per capita and sectoral capital intensity by the capital stock per employee. I construct state and sectoral capital stocks using a perpetual inventory method.[26] Denote capital stocks by K, capital expenditures by I and the geometric decay rate by δ. State and sectoral capital stocks evolve according to

$$K_{i,t} = (1-\delta)\, K_{i,t-1} + I_{i,t} \tag{2.3a}$$

$$K_{s,t} = (1-\delta)\, K_{s,t-1} + I_{s,t} \tag{2.3b}$$

[26]The US Census' Quarterly Financial Report directly measures capital stocks for 3-digit NAICS sectors, so I do not need a perpetual inventory approach for sectoral capital intensities for the robustness check with disaggregated energy intensities in the Appendix.

Harberger (1978) notes that when capital growth $I_t/K_{t-1} - \delta$ equals output growth g, the initial capital stock can be calculated as

$$K_{t-1} = I_t / (\delta + g) \tag{2.4}$$

I set $\delta = 0.035$ and $g = 0.03$. My data on state capital expenditures start in 1987; the data on sectoral expenditures in 1997, because of the transition from the SIC to NAICS classification.

2.A.2. Supplementary tables and robustness checks

Table 2.9 presents the first-stage regression results for specification (2.2), which are discussed in the main text. The results indicate that energy prices are lower in energy-abundant states. The first-stage coefficients should be interpreted with care, as they are overproportionally driven by states with large energy-intensive sectors. The coefficient on the coal interaction term implies that a one standard deviation increase in coal abundance is associated with a decrease in electricity prices of 0.45 standard deviations.

Table 2.10 shows the regression coefficients of (2.1) when I normalize energy- and capital abundance by land area instead of population. The coefficients on all interaction terms for energy and capital are smaller in absolute value than in the main specification. The results do not suggest that energy and capital endowments attract migration. The coefficients for coal and hydro decrease more than for natural gas, as area normalizations introduce more noise into the coal and hydro abundance than in the natural gas abundance measure. The variation in population density between states with the largest absolute coal[27] and hydro[28] endowments is larger than between states with the largest natural gas[29] endowments.

Table 2.12 lists the results for regression equation (2.1) when I decompose fuel intensity into natural gas intensity, distillate fuel oil intensity and residual fuel oil intensity. The correlation between the natural gas and oil interaction terms is much lower than in the main specification, as can be seen in Table 2.11. Natural gas intensity and electricity intensity are perfectly correlated however. The negative coefficient on the oil interac-

[27] Montana, Illinois, Wyoming, Kentucky, West Virginia
[28] Washington, California, Oregon, New York, Alabama
[29] Texas, Wyoming, Oklahoma, Colorado, New Mexico

Table 2.7: State Characteristics

Variable	Definition	Source
Skilled labour abundance	Fraction of population over 25 with a Bachelor's degree or higher	US Census Bureau
Capital expenditures	Capital expenditures in all manufacturing sectors	Annual Survey of Manufactures
Capital abundance	Capital stock per capita	Eqns (2.3a) and (2.4)
Coal abundance	Estimated recoverable coal reserves per capita	Energy Information Administration
Natural gas abundance	Dry natural gas reserves per capita	Energy Information Administration
Oil abundance	Crude oil reserves per capita [a]	Energy Information Administration
Hydro abundance	Summer capacity hydroelectric generation per capita	Energy Information Administration
Electricity price	Electricity price in the industrial sector	State Energy Data System, EIA
Natural gas price	Natural gas price in the industrial sector	State Energy Data System, EIA
Distillate fuel oil price	Distillate fuel oil price in the industrial sector	State Energy Data System, EIA
Residual fuel oil price	Residual fuel oil price in the industrial sector	State Energy Data System, EIA
Market potential	$\sum_{i'}\left(\frac{\text{population in state } i' \text{ in } 100.000}{\max\left(\text{distance between states } i \text{ and } i' \text{ in miles}, {}^{b}100\right)}\right)$	US Census Bureau
Electricity deregulation	1 if electricity restructuring legislation enabled by July 2000, 0 o.w.	(EIA, 2000)
Natural gas price cap	1 if price caps implemented in 2000 or earlier, 0 otherwise	(Hlasny, 2008)
Natural gas cost incentive	1 if cost incentive measures implemented in 2000 or earlier, 0 otherwise	(Hlasny, 2008)
Heating degree days	Total heating degree days weighted by population	NOAA

All state characteristics are truncated at the 95th percentile and divided by the yearly standard deviation. I therefore provide no units of measurement. Except for the electricity and natural gas deregulation indicators, the dimension of all state characteristics is i, t. Monetary values are adjusted by the consumer price index from the Bureau of Economic Analysis. [a] Does not include Federal Offshore Reserves [b] I follow Harris (1954). Distances are for the quickest route between the largest cities (2000) in both states (Source: www.mileage-charts.com). Distances to and from Burlington, VT and Honolulu, HI are from Google Maps.

Table 2.8: Sector Characteristics

Variable	Dimension	Definition	Source
Value added	i, s, t	Value added	Annual Survey of Manufactures
Employment	i, s, t	Number of employees	Annual Survey of Manufactures
Skilled labour intensity	s, t	Nationwide average wage per employee	Annual Survey of Manufactures
Capital expenditures	s, t	Capital expenditures in all manufacturing sectors	Annual Survey of Manufactures
Capital intensity	s, t	Capital stock per employee	Eqns (2.3b) and (2.4)
Electricity intensity	s, t	Quantity of electricity purchased per employee	Annual Survey of Manufactures
Fuel intensity	s, t	Cost of purchased fuels per employee	Annual Survey of Manufactures
Forward linkages	s^a	Row sum of the Ghosh inverse	Bureau of Economic Analysis
Backward linkages	s^a	Column sum of the Leontief inverse	Bureau of Economic Analysis
Scale economies	s, t	Nationwide average establishment size	County Business Patterns

Sector characteristics are equal to one (zero) if the sector characteristic is higher (lower) than the yearly average across sectors. [a] The input-output matrix is only available for 2002. I equate forward and backward linkages to the 2002 value for the whole sample period.

tion term in the main specification disappears, though the oil interaction terms are not significant when I control for natural gas abundance and intensity.

Table 2.13 presents the results when I use the log of employment, rather than the log of value added, as dependent variable. All interaction terms have slightly smaller coefficients than in the main specification, though the significance levels are unchanged. A possible explanation is that factor prices are lower in states in which the factor is abundant. For a given output level, capital-intensive industries then have lower costs, and hence a higher value added, in capital-abundant states. By contrast, the number of employees per unit of physical output is more likely to be constant across capital-abundant and capital-scarce states.

Table 2.14 contains the coefficient estimates when energy- and capital intensity are defined as energy expenditures and capital stock per dollar of value added, respectively. Like in Table 2.13, the coefficients are slightly lower in absolute value. The distribution of energy expenditures per employee is more skewed than that of energy expenditures per dollar of value added. Therefore, the average energy-intensity using the alternative metric is lower than the average intensity using the main metric, and more industries classify as energy intensive than in the main specification. Since energy endowments are less important in the location decision of the marginal energy-intensive industries, the energy interaction terms lose some explanatory power.

Table 2.15 shows the results when I do not discretize the sector intensities, to verify whether we lose information as a result of the discretization. The first energy interaction term, as well as the capital and skill interaction terms, are slightly larger than in the main specification. The electricity intensity $\times$ hydro abundance coefficient is no longer significant, suggesting that hydro abundance does not play as large of a role in the location of the most electricity-intensive industries. The correlation coefficient between electricity intensity $\times$ natural gas abundance and fuel intensity $\times$ natural gas abundance is higher than in the main specification (0.84 instead of 0.81). Due to increased multicollinearity, the coefficients on the natural gas interaction terms in column (3) diverge.

Table 2.9: First-stage coefficient estimates for instrumental variables regressions

Dependent variable	(1) electricity intensity × electricity price	(2) fuel intensity × natural gas price	(3) fuel intensity × distillate fuel oil price	(4) fuel intensity × residual fuel oil price
electricity intensity × electricity deregulation	0.77 (0.050)***			
fuel intensity × natural gas price cap		0.35 (0.066)***		
fuel intensity × natural gas cost incentive		0.05 (0.052)		
fuel intensity × heating degree days		0.08 (0.029)***	0.43 (0.028)***	-0.11 (0.024)***
electricity intensity × coal abundance	-0.45 (0.027)***			
electricity intensity × hydro abundance	-0.13 (0.020)***			
electricity intensity × natural gas abundance	-0.20 (0.035)***			
fuel intensity × natural gas abundance		-0.38 (0.059)***		
fuel intensity × oil abundance			-0.11 (0.048)**	-0.40 (0.046)***
Number of observations	18423	18908	18908	18908
Number of clusters	2971	3071	3071	3071
Adjusted R^2	0.94	0.98	1.00	0.97

This table reports first-stage coefficient estimates for the GMM-IV regressions in Table 2.6. Capital, skilled labour and new economic geography interaction terms as well as two-way fixed effects (state-year and sector-year) are included. Error terms are clustered by sector-state pair. Standard errors in parentheses. Asterisks denote significance at the 10% (*), 5% (**) and 1% (***) level. Sector characteristics are equal to one if they are larger than average, and zero if they are smaller than average. State characteristics are divided by their standard deviation.

Table 2.10: Location of 4-digit NAICS industries with area-normalized energy and capital endowments

	(1)	(2)	(3)
electricity intensity × coal abundance	0.07 (0.041)*		0.07 (0.040)*
electricity intensity × natural gas abundance	0.16 (0.052)***		−0.00 (0.053)
electricity intensity × hydro abundance	−0.01 (0.039)		0.00 (0.038)
fuel intensity × natural gas abundance		0.33 (0.057)***	0.34 (0.063)***
fuel intensity × oil abundance		−0.14 (0.050)***	−0.13 (0.050)***
capital intensity × capital abundance	−0.01 (0.040)	0.01 (0.039)	0.00 (0.040)
skill intensity × skill abundance	0.04 (0.031)	0.05 (0.031)	0.04 (0.031)
forward linkages × market potential	0.00 (0.035)	0.00 (0.035)	0.00 (0.035)
backward linkages × market potential	−0.13 (0.036)***	−0.13 (0.035)***	−0.13 (0.036)***
scale economies × market potential	−0.08 (0.039)**	−0.08 (0.038)**	−0.09 (0.039)**
Number of observations	18423	18908	18423
Number of clusters	2971	3071	2971
Adjusted R^2	0.57	0.57	0.57

This table reports coefficient estimates for regression equation (2.1). Energy and capital endowments are measured as reserves or stock per square mile of land area, truncated at the 95th percentile and divided by the yearly standard deviation. Other variable definitions are the same as in Tables 2.7 and 2.8. Two-way fixed effects (state-year and sector-year) are included. Error terms are clustered by sector-state pair. Standard errors in parentheses. Asterisks denote significance at the 10% (*), 5% (**) and 1% (***) level. Sector characteristics are equal to one if they are larger than average, and zero if they are smaller than average. State characteristics are divided by their standard deviation.

Table 2.11: Sample correlation coefficients for energy interaction terms with disaggregated energy intensities

	(1)	(2)	(3)	(4)	(5)	(6)
(1) electricity intensity × coal abundance	1					
(2) electricity intensity × natural gas abundance	0.18	1				
(3) electricity intensity × hydro abundance	0.09	−0.02	1			
(4) natural gas intensity × natural gas abundance	0.18	1	−0.02	1		
(5) distillate fuel oil intensity × oil abundance	0.28	0.45	0.01	0.45	1	
(6) residual fuel oil intensity × oil abundance	0.32	0.49	0.03	0.49	0.72	1

Table 2.12: Location of 3-digit NAICS industries with disaggregated energy intensities

	(1)	(2)	(3)	(4)
electricity intensity × coal abundance	0.25 (0.084)***			0.23 (0.086)***
electricity intensity × natural gas abundance	0.23 (0.088)***			
electricity intensity × hydro abundance	0.03 (0.064)			0.03 (0.065)
natural gas intensity × natural gas abundance		0.33 (0.083)***		0.24 (0.085)***
distillate fuel oil intensity × oil abundance			0.23 (0.086)***	0.10 (0.084)
residual fuel oil intensity × oil abundance			0.17 (0.097)*	0.10 (0.091)
capital intensity × capital abundance	0.06 (0.059)	0.06 (0.059)	0.04 (0.058)	0.07 (0.059)
skill intensity × skill abundance	−0.01 (0.047)	−0.00 (0.047)	−0.01 (0.046)	−0.00 (0.047)
forward linkages × market potential	0.09 (0.054)*	0.09 (0.052)*	0.08 (0.052)	0.10 (0.054)*
backward linkages × market potential	−0.11 (0.065)*	−0.11 (0.063)*	−0.11 (0.063)*	−0.11 (0.064)*
scale economies × market potential	0.12 (0.063)**	0.14 (0.062)**	0.10 (0.062)*	0.11 (0.064)*
Number of observations	7080	7335	7335	7080
Number of clusters	931	968	968	931
Adjusted R^2	0.73	0.73	0.73	0.73

This table reports coefficient estimates for regression equation (2.1). Energy intensities, defined as energy use per employee, are from the Manufacturing Energy Consumption Survey. These data are only available for 2002 and 2006. I use the 2002 values for 2001-2004 and the 2006 values for 2005-2009. Capital intensity is defined as net property, plant and equipment per employee from the Quarterly Financial Report of the US Census. I constructed yearly data by taking averages across quarters. Other variable definitions are the same as in Tables 2.7 and 2.8. Two-way fixed effects (state-year and sector-year) are included. Error terms are clustered by sector-state pair. Standard errors in parentheses. Asterisks denote significance at the 10% (*), 5% (**) and 1% (***) level. Sector characteristics are equal to one if they are larger than average, and zero if they are smaller than average. State characteristics are divided by their standard deviation.

Table 2.13: Location of 4-digit NAICS industries with log of employment as dependent variable

	(1)	(2)	(3)
electricity intensity × coal abundance	0.17		0.20
	(0.046)***		(0.049)***
electricity intensity × natural gas abundance	0.14		−0.08
	(0.048)***		(0.063)
electricity intensity × hydro abundance	0.14		0.14
	(0.039)***		(0.039)***
fuel intensity × natural gas abundance		0.26	0.37
		(0.058)***	(0.075)***
fuel intensity × oil abundance		−0.06	−0.16
		(0.083)	(0.089)*
capital intensity × capital abundance	0.02	0.02	0.02
	(0.036)	(0.036)	(0.036)
skill intensity × skill abundance	0.04	0.05	0.04
	(0.025)	(0.025)*	(0.025)*
forward linkages × market potential	0.06	0.03	0.06
	(0.030)**	(0.029)	(0.030)**
backward linkages × market potential	−0.10	−0.10	−0.10
	(0.030)***	(0.030)***	(0.030)***
scale economies × market potential	−0.05	−0.05	−0.05
	(0.032)	(0.032)	(0.032)
Number of observations	19373	19896	19373
Number of clusters	3109	3210	3109
Adjusted R^2	0.52	0.51	0.52

This table reports coefficient estimates for regression equation (2.1) with log of employment, rather than log of value added, as dependent variable. Two-way fixed effects (state-year and sector-year) are included. Error terms are clustered by sector-state pair. Standard errors in parentheses. Asterisks denote significance at the 10% (*), 5% (**) and 1% (***) level. Sector characteristics are equal to one if they are larger than average, and zero if they are smaller than average. State characteristics are divided by their standard deviation.

Table 2.14: Location of 4-digit NAICS industries with alternative energy and capital intensities

	(1)	(2)	(3)
electricity intensity × coal abundance	0.21		0.22
	(0.055)***		(0.057)***
electricity intensity × natural gas abundance	0.17		−0.05
	(0.056)***		(0.106)
electricity intensity × hydro abundance	0.13		0.13
	(0.045)***		(0.045)***
fuel intensity × natural gas abundance		0.24	0.34
		(0.074)***	(0.118)***
fuel intensity × oil abundance		−0.02	−0.11
		(0.101)	(0.104)
capital intensity × capital abundance	0.16	0.16	0.16
	(0.032)***	(0.032)***	(0.032)***
skill intensity × skill abundance	0.04	0.05	0.04
	(0.031)	(0.030)	(0.031)
forward linkages × market potential	0.03	−0.00	0.03
	(0.037)	(0.035)	(0.037)
backward linkages × market potential	−0.12	−0.13	−0.12
	(0.035)***	(0.035)***	(0.035)***
scale economies × market potential	−0.08	−0.08	−0.08
	(0.039)**	(0.039)**	(0.039)**
Number of observations	18423	18908	18423
Number of clusters	2971	3071	2971
Adjusted R^2	0.57	0.57	0.57

This table reports coefficient estimates for regression equation (2.1). Energy and capital intensities are measured as expenditures or stock per dollar of value added. Other variable definitions are the same as in Tables 2.7 and 2.8. Two-way fixed effects (state-year and sector-year) are included. Error terms are clustered by sector-state pair. Standard errors in parentheses. Asterisks denote significance at the 10% (*), 5% (**) and 1% (***) level. Sector characteristics are equal to one if they are larger than average, and zero if they are smaller than average. State characteristics are divided by their standard deviation.

Table 2.15: Location of 4-digit NAICS industries with continuous sector characteristics

	(1)	(2)	(3)
electricity intensity × coal abundance	0.27		0.28
	(0.044)***		(0.048)***
electricity intensity × natural gas abundance	0.30		−0.26
	(0.076)***		(0.182)
electricity intensity × hydro abundance	0.07		0.07
	(0.045)		(0.045)
fuel intensity × natural gas abundance		0.36	0.67
		(0.099)***	(0.194)***
fuel intensity × oil abundance		−0.01	−0.15
		(0.129)	(0.137)
capital intensity × capital abundance	0.11	0.09	0.10
	(0.046)**	(0.047)**	(0.046)**
skill intensity × skill abundance	0.10	0.08	0.10
	(0.038)**	(0.038)**	(0.038)***
forward linkages × market potential	0.02	−0.00	0.01
	(0.042)	(0.040)	(0.042)
backward linkages × market potential	−0.17	−0.18	−0.17
	(0.044)***	(0.043)***	(0.044)***
scale economies × market potential	−0.05	−0.05	−0.05
	(0.052)	(0.052)	(0.052)
Number of observations	18423	18908	18423
Number of clusters	2971	3071	2971
Adjusted R^2	0.57	0.57	0.57

This table reports coefficient estimates for regression equation (2.1). Sector characteristics are truncated at the 95th percentile and normalized so that the yearly mean and standard deviation across sectors is 1/2. Definitions of state characteristics are listed in Table 2.7; sources of sector characteristics are the same as in Table 2.8. Two-way fixed effects (state-year and sector-year) are included. Error terms are clustered by sector-state pair. Standard errors in parentheses. Asterisks denote significance at the 10% (*), 5% (**) and 1% (***) level. State characteristics are divided by their standard deviation.

Chapter 3

Brown backstops versus the green paradox[30]

3.1. Introduction

Well-intended climate policies may have perverse effects. Climate policies typically become stricter over time. Fossil fuel owners, deciding when to sell their scarce resources, may respond by speeding up extraction. This side effect can occur when fossil fuel reserves are limited and cheap to exploit: a reasonable characterization for conventional oil and natural gas, but much less for other important energy sources such as coal and unconventional oil. In this paper we ask whether climate policy has unintended consequences when there are two types of fossil fuels: one dirty and scarce, the other even dirtier and abundant.

Policies that reduce future dependence on fossil fuels might encourage suppliers, anticipating a future drop in demand, to bring forward the extraction of their resources. When present emissions are more harmful than future emissions, gradually increasing carbon taxes can be counterproductive: a green paradox (Sinn, 2008a). Developing a carbon-free substitute for fossil fuels can cause a similar effect (Strand, 2007; Hoel, 2011). Cost reductions for the substitute decrease the scarcity value of fossil fuels, and thereby increase fossil fuel supply in all periods before exhaustion.[31]

The crucial feature that drives the above mechanism is the exhaustibility of the resource. This causes the tradeoff between current and future supply, and thus the

[30] An earlier version of this article circulates as Michielsen (2011).

[31] The green paradox may vanish when the substitute has an upward-sloping supply curve (Gerlagh, 2011). Van der Ploeg and Withagen (2012a) find that the green paradox occurs for clean but expensive backstops (such as solar or wind), but not when the backstop is sufficiently cheap relative to emissions damages, as it is then attractive to leave part of the oil in the ground.

effect of (expected) future policies on current supply and emissions. If the resource is fully abundant, resource owners supply the myopically optimal quantity in each period and the link between current and future markets is severed. Exhaustibility is a fair assumption for conventional oil and natural gas, which will be depleted in 50 to 70 years at current consumption rates.[32] Coal and unconventional oil are much more abundant however. Coal reserves are sufficient to last another 250 years, and tar sand deposits in Alberta are estimated at 1800 bln barrels.[33] The supply of these resources is primarily driven by costs rather than scarcity rents. Anticipated carbon taxes cause coal mines to shut down in the future, but do not increase near-term supply.

Coal and unconventional oil are significant from an economic and a climate change point of view. Coal satisfies a third of global energy demand and accounts for almost half of energy-related CO2 emissions,[34] outranking petroleum in emission intensity by 30-40%. The IEA expects coal supply to increase by 60% in 2035 under business-as-usual policies;[35] twice as much as the projected increase in oil supply. Supply of unconventional oil, which is 20% more emission-intensive than petroleum (Charpentier et al., 2009), may increase fivefold to 11 mln barrels per day in 2035. These numbers suggest that in order to keep climate change within tolerable limits, it is imperative that coal and uncoventional oil reserves remain largely unexploited (Gerlagh, 2011). A comprehensive assessment of the effectiveness of climate policies should take into account these dirty substitutes and their unique characteristics.[36]

In this paper, we develop a simple model with two time periods. We do not derive optimal policies, but present a descriptive analysis of the effect of future climate policies on emissions. We generalize assumptions in previous research along two important dimensions. Firstly, the model contains three energy types: a dirty exhaustible resource (e.g. oil), an even dirtier substitute (coal) and a clean substitute (solar). Secondly, we assume types to be imperfect substitutes for one another. Previous theoretical studies often assume perfect substitution, which is unrealistic. We model climate policy as a carbon tax or a decrease in the cost of the clean substitute. We calculate intertemporal

[32] BP (2010, p. 6, p. 12)
[33] Alberta's Energy Reserves 2010 and Supply/Demand Outlook 2011-2020, p.5
[34] International Energy Statistics, Energy Information Administration
[35] IEA (2010b, p. 201)
[36] Van der Ploeg and Withagen (2012b) show that rising carbon taxes may not cause a green paradox when coal, rather than renewables, is the primary alternative for oil.

carbon leakage as the increase in present emissions over the decrease in future emissions.

By virtue of the abundance of their resource, coal owners do not trade off present and future extraction. When faced with a demand reduction in the future, they will therefore not increase supply today. Oil emissions may leak away to the present, but the increase in current oil supply reduces demand for dirtier coal. Carbon taxes can cause negative leakage when the substitutability between oil and coal differs between periods. We may call this a 'strong green orthodox' (Grafton et al., 2012). Moreover, since carbon taxes decrease the price of oil relative to coal, a future tax delays rather than accelerates oil extraction when oil and coal are good substitutes in the future. Reducing the future cost of solar decreases present emissions when oil and coal are good substitutes or if the emission intensity of coal is high.

Our contribution is twofold. Firstly, we offer a general theoretical framework that can make more accurate predictions than models that include only one or two energy types or assume perfect substitutability. The presence of an abundant dirty substitute reduces intertemporal leakage directly and indirectly, and may even cause negative leakage rates. By making more specific assumptions, we can obtain similar findings as in other papers on the green paradox. Secondly, our model is well suited for empirical calibration. For carbon taxes, intertemporal leakage rates are negative or less than 5%. For reductions in the future cost of renewables, leakage is between 2-13% for biofuels and 0-2% for solar or wind electricity. Biofuels are a close substitute for oil, the most emission-intensive scarce fossil fuel, and hence more prone to intertemporal leakage than renewable electricity, which primarily competes with coal.

Though we focus on intertemporal leakage, our framework can also be used to analyze spatial carbon leakage, by relabeling the two time periods as two countries and setting the interest rate to zero. Calibrating a spatial version of the model, we find leakage rates ranging from negative to 40%, comparable to estimates from computable general equilibrium models (Di Maria and van der Werf, 2012). These findings suggest that the green paradox is a small concern relative to spatial carbon leakage.

The rest of this paper is organized as follows. Section 3.2 outlines the model. Section 3.3 analyzes intertemporal and spatial leakage when carbon emissions are taxed in the future. Section 3.4 studies the impact of reductions in the future cost of a clean substitute. We calibrate the models in section 3.5. Section 3.6 discusses implications of the model for spatial carbon leakage, and calibrates a spatial version of the model. Section 3.7 concludes. All proofs are relegated to the Appendix.

3.2. Model

Consider a model with three types of energy: an exhaustible resource, a dirty backstop and a clean backstop. The backstops are inexhaustible, supplied competitively and have constant marginal costs.[37] Though the word 'backstop' is sometimes used to denote a perfect substitute for an exhaustible resource, we explicitly allow for imperfect substitutability. The exhaustible resource is supplied competitively by a group of energy exporters and costless to extract. For the energy exporters, it is always optimal to fully exhaust the fossil resource stock S. An energy-importing country derives utility from consuming energy. Denote the exhaustible resource, the dirty and the clean backstop with superscripts R, D and C respectively. Demand functions are given by

$$d^i(p^i, p^{-i}), \ i \in \{R, D, C\} \tag{3.1}$$

where p^i is the consumer price of resource i. Throughout the paper, we write shorthand d^i for (6.8). Partial derivatives of d^i are indicated by a subscript of the corresponding type. We make the following assumptions about energy demand

$$d^i_i < 0, \ d^i_j \geq 0 \tag{3.A1}$$

$$\left|d^i_i\right| > \left|d^i_j\right|, \ i \neq j \tag{3.A2}$$

$$d^i_j = d^j_i \tag{3.A3}$$

Energy types are imperfect substitutes for one another: demand for each type is non-decreasing in the price of other types (3.A1) and own-price effects are larger than cross-price effects (6.10). Cross-price effects are symmetric (3.A3). Assumption (3.A3) is not necessary for many of our results; we explicate any invocation of (3.A3) in the proposition texts. The assumptions hold if the relative budget shares of the three energy

[37] An upward-sloping supply curve for the clean backstop reduces intertemporal carbon leakage (Gerlagh, 2011).

types do not depend on available income.

Consumption of the exhaustible resource and the dirty backstop generates a constant amount of emissions. The dirty backstop is more emission intensive than the exhaustible resource

$$e = \zeta^R d^R + \zeta^D d^D, 0 < \zeta^R < \zeta^D$$

The model consists of two periods. All variables corresponding to the second period are denoted by capitals. We allow for emissions in the first period to be more harmful than emissions in the second period. Total emission damages are

$$\Sigma = e + \beta E,\ \ \beta \leq 1 \tag{3.3}$$

When only cumulative emissions matter, β is equal to one. When society and ecology can adapt more easily to slow rather than rapid temperature increases (Hoel and Kverndokk, 1996; Gerlagh, 2011), near-term emissions have a higher weight ($\beta < 1$). The green paradox entails a positive relation between the stringency of future climate policy and emissions (Sinn, 2008b). Following Gerlagh (2011), we differentiate between a weak green paradox (future climate policy increases present emissions) and a strong green paradox (emission damages increase).

Definition 3.1. *Denote the stringency of second-period climate policy by* Θ*. The weak green paradox occurs if*

$$\frac{\mathrm{d}e}{\mathrm{d}\Theta} > 0$$

The strong green paradox occurs if

$$\frac{\mathrm{d}\Sigma}{\mathrm{d}\Theta} > 0$$

Analogous to the literature on (spatial) carbon leakage, we define the intertemporal carbon leakage of a future climate policy as the share of period 2 emission reductions that 'leaks' away to the first period.

Definition 3.2. *The leakage* λ *of an increase in the stringency of second-period policy* Θ *is the increase in period 1 emissions over the decrease in period 2 emissions.*

$$\lambda \equiv -\ \frac{\mathrm{d}e}{\mathrm{d}\Theta}\ /\ \frac{\mathrm{d}E}{\mathrm{d}\Theta}$$

Both green paradoxes are related to the intertemporal leakage rate λ in a straightforward way. As intertemporal leakage is positive if and only if the future climate policy increases present emissions, the weak green paradox is equivalent to $\lambda > 0$. The strong green paradox occurs if the leakage rate exceeds the emission discount factor ($\lambda > \beta$).

Exhaustible resource owners discount future revenues at rate r. In equilibrium, they are indifferent between extracting now and in the future. Letting Π^R be the second-period producer price of the exhaustible resource, the Hotelling condition reads

$$p^R = \frac{1}{1+r}\Pi^R\left(P^R, \Theta\right) \tag{3.4}$$

We discuss carbon taxes (section 3.3) and investment in green technologies (section 3.4) in turn.

3.3. Emission taxes

Regulators who want to reduce carbon emissions may not be able to do so immediately. Swift implementation of climate policies is often impeded by political and technological considerations. Announcing carbon taxes or caps in advance reduces compliance costs: it gives firms the opportunity to purchase abatement equipment and adjust their production processes, and allows consumers to make informed decisions about durable good purchases (Di Maria et al., 2008). The European Commission notes that "a sufficient carbon price and long-term predictability are necessary"[38] in order to meet the 80-95% EU emission reduction target in 2050.[39]

Carbon emissions are taxed at a constant rate T in the second period. The tax may also be interpreted as a willingness to pay to reduce emissions (Hoel, 2010). Exhaustible resource owners discount future receipts net of the tax at the interest rate:

$$p^R = \frac{1}{1+r}\left(P^R - T\zeta^R\right)$$

A second-period carbon tax only affects first-period variables through the exhaustible

[38] A Roadmap for moving to a competitive low carbon economy in 2050, p.7, European Commission COM(2011) 112

[39] Announcements should of course be credible. For discussions on credibility issues in climate policy, see (Helm et al., 2003; Golombek et al., 2010).

resource price. The change in first-period emissions is

$$\frac{\mathrm{d}e}{\mathrm{d}T} = \frac{\partial e}{\partial p^R}\frac{\mathrm{d}p^R}{\mathrm{d}T} \tag{3.5}$$

We discuss the two components of the right-hand side in turn. The carbon tax increases the period 2 producer price of the exhaustible resource and, by (3.4), the period 1 price if and only if the tax increases period 2 exhaustible resource demand at the initial producer price.

$$\frac{\mathrm{d}p^R}{\mathrm{d}T} = \frac{1}{1+r}\frac{\mathrm{d}\Pi^R}{\mathrm{d}T} \gtreqless 0 \Leftrightarrow \frac{\partial D^R}{\partial T} \gtreqless 0 \tag{3.6}$$

Holding the producer price constant, the carbon tax directly reduces exhaustible resource demand in the second period by $-\zeta^R D^R_R$. The tax has an even stronger effect on the future price of the dirty backstop by virtue of its higher emission intensity however. This induces substitution from the dirty backstop to the exhaustible resource, increasing future exhaustible resource demand by $\zeta^D D^R_D$. The period 1 exhaustible resource price goes up if the net effect of the tax on period 2 exhaustible resource demand

$$\frac{\partial D^R}{\partial T} = \zeta^R D^R_R + \zeta^D D^R_D \tag{3.7}$$

is positive, i.e. if the substitutability between the dirty backstop and the exhaustible resource is high in period 2 and if the emission intensity of the dirty backstop is high. Conversely, a period 2 carbon tax decreases exhaustible resource prices if the dirty backstop and the exhaustible resource are poor substitutes in period 2 and if the dirty backstop has a low emission-intensity.

The effect of exhaustible resource prices on period 1 emissions is similar. An increase in the period 1 exhaustible resource price directly reduces emissions by $-\zeta^R d^R_R$. Higher exhaustible resource prices also encourage substitution towards the dirty backstop, increasing emissions by $\zeta^D d^D_R$. The net change in emissions

$$\frac{\partial e}{\partial p^R} = \zeta^R d^R_R + \zeta^D d^D_R \tag{3.8}$$

is positive if the dirty backstop and the exhaustible resource are good substitutes in the first period and if the emission intensity of the dirty backstop is high. On the other hand, higher exhaustible resource prices decrease first-period emissions when the substitutability between the exhaustible resource and the dirty backstop is low in the first period and when the dirty backstop is not very emission intensive. Proposition 6.3 gives the condition for positive leakage.

Proposition 3.1 (weak green paradox)**.** *Following a carbon tax increase in period 2,* $\lambda \gtreqless 0$ *iff*

$$\left(\zeta^R d_R^R + \zeta^D d_R^D\right)\left(\zeta^R D_R^R + \zeta^D D_D^R\right) \gtreqless 0 \tag{3.9}$$

A weak green paradox is less likely if the substitutability between the exhaustible resource and the dirty backstop is different in the two periods. Table 3.1 summarizes whether the weak green paradox occurs for different values of d_R^D and D_D^R and how these cases relate to previous research.

When the exhaustible resource and the dirty backstop are poor substitutes in both periods (d_R^D and D_D^R are both low), the future tax reduces exhaustible resource prices and increases emissions in the first period. This is the classic green paradox result when exhaustible resource owners anticipate a future carbon tax (Hoel, 2010). When substitutability between the exhaustible resource and the dirty backstop is low in the first period but high in the second (d_R^D is low, while D_D^R is high), the tax increases exhaustible resource prices and reduces emissions in both periods. Exhaustible resource owners delay extraction in response to the tax, as the tax puts them at a comparative advantage vis a vis the dirty backstop in the future. Since the dirty backstop is a poor substitute for the exhaustible resource in the short term, the decline in period 1 exhaustible resource supply does not cause a surge in dirty backstop demand. Our model provides a theoretical framework for the numerical findings of Persson et al. (2007). They show that OPEC countries may benefit rather than lose from strict climate policies, because the price of synthetic substitutes for petroleum-based fuels (e.g. diesel from coal) goes up faster than the price of oil.

When substitutability is high in the first period but low in the second (d_R^D is high, but D_D^R is low), the second-period tax reduces exhaustible resource prices and emissions go down in both periods. As the substitutability between the exhaustible resource and the dirty backstop is low in the second period, exhaustible resource prices decrease. This makes the exhaustible resource an attractive alternative to the dirty backstop in the first period. Lastly, suppose that exhaustible resource and the dirty backstop are good substitutes in both periods (d_R^D and D_D^R are both high). The tax then increases exhaustible

Table 3.1: Occurrence of the weak green paradox for different values of d_R^D and D_D^R

d_R^D	D_R^D	$\frac{\mathrm{d}p^R}{\mathrm{d}T}$	$\frac{\partial e}{\partial p^R}$	weak GP?	Related articles
low	low	-	-	yes	Sinn (2008a); Hoel (2010)
low	high	+	-	no	Persson et al. (2007)
high	low	-	+	no	
high	high	+	+	yes	Smulders and van der Werf (2008) Di Maria et al. (2008)

resource prices and increases emissions in the first period, as the dirty backstop is used more intensively early on. This result connects to work of Smulders and van der Werf (2008) and Di Maria et al. (2008), who analyze how an anticipated cap on the flow of emissions affects the order of extraction when there is a high- and a low-carbon fuel. The cap makes the low-carbon fuel more valuable and increases the use of the high-carbon fuel in the period before the constraint becomes active.

Proposition 6.4 describes the effects of a period 2 tax on period 2 emissions and emission damages.

Proposition 3.2 (strong green paradox)**.** *Following a carbon tax increase in period 2*
(i) D^D decreases
(ii) under (3.A3)*, E decreases*
(iii) under (3.A3)*, $\lambda \leq 1$*
(iv) $\lambda > \beta$ iff

$$-\frac{\zeta^R D_R^R + \zeta^D D_D^R}{d_R^R + (1+r)\, D_R^R}\, \Xi + \beta\zeta^D \left(\zeta^R D_R^D + \zeta^D D_D^D\right) > 0, \;\; \textit{where}$$
$$\Xi = (1-\beta)\, \zeta^R d_R^R + \zeta^D \left(d_R^D + \beta\,(1+r)\, D_R^D\right) \tag{3.10}$$

(v) λ decreases in $|D_D^D|$

A higher own-price effect of the dirty backstop causes the tax to more sharply reduce period 2 dirty backstop use, and therefore reduces leakage. The effect of the own- and cross-price effects of the exhaustible resource on the intertemporal leakage rate cannot be signed because the effect of the tax on exhaustible resource extraction is ambiguous.[40] Although the tax increases future demand for the clean backstop, clean backstop prices, quantities and elasticities do not appear in the conditions for $\lambda > 0$ and $\lambda > \beta$. The clean backstop does not generate emissions, so d^C does not enter into either e or Σ. Furthermore, the tax does not affect the price of the clean backstop, so $\frac{\mathrm{d}e}{\mathrm{d}T}$ and $\frac{\mathrm{d}\Sigma}{\mathrm{d}T}$ do not contain any derivatives with respect to p^C. The impact of the clean backstop on intertemporal leakage is implicit in the demand functions for the exhaustible resource and the dirty backstop.

Interpreting (3.10) is not straightforward, but we can calibrate λ as estimates of all parameters in (3.10) are available. Own- and cross-price effects can be rewritten as $d^i_j = \eta^i_j d^i / p^j$, where η^i_j is the elasticity of demand for type i with respect to the price of type j. We estimate the magnitude of intertemporal carbon leakage in section 3.5.

3.4. A cheaper clean backstop

In addition to implementing a carbon tax, climate-conscious policymakers may opt to reduce emissions by stimulating the development of clean alternatives to fossil fuels. To model such a policy, we analyze the effect of a reduction in the period 2 price of the clean backstop P^C on emissions. The development of alternative energy sources requires resources to be committed well before the new technologies can be put to use, so exhaustible resource owners anticipate the lower period 2 clean backstop prices when deciding on the intertemporal extraction pattern. A lower P^C reduces exhaustible resource demand in period 2, and thus decreases the right hand side of (3.4). For exhaustible resource owners to remain indifferent between extracting in either period, period 1 extraction d^R must go up. This is the classic green paradox result (Strand, 2007; Hoel, 2011). The improved technology also reduces emissions from the dirty backstop however. In the next Propositions, we show how the occurrence of the weak and the strong green paradox depend on the emission intensities and the substitutability between energy types.

Proposition 3.3 (weak green paradox)**.** *Assume $D^R_C > 0$. When the clean backstop*

[40] Albeit through a different mechanism (intertemporal substitution in consumption rather than substitution between energy types), (Eichner and Pethig, 2011) also find that a future emission constraint need not cause a green paradox.

becomes cheaper in period 2, $\lambda \gtreqless 0$ *iff*

$$\zeta^R d_R^R + \zeta^D d_R^D \lesseqgtr 0 \tag{3.11}$$

As opposed to the case of a future carbon tax, exhaustible resource owners always bring forward extraction when clean alternatives become cheaper in the future. The lower exhaustible resource prices also causes a drop in period 1 demand for the dirty backstop. The occurrence of the weak green paradox hinges on whether the increase in exhaustible resource-related emissions outweighs the decrease in dirty backstop-related emissions (3.11). This is more likely if the relative emission intensity of the exhaustible resource is high and if the substitutability between the exhaustible resource and the dirty backstop is low. All first-period effects are proportional to the change in the period 1 exhaustible resource price $\frac{\mathrm{d}p^R}{\mathrm{d}P^C}$. Because period 2 parameters only affect period 1 emissions through this term, the condition for the weak green paradox consists solely of period 1 parameters.

Proposition 3.4 (strong green paradox)**.** *When the clean backstop becomes cheaper in period 2,*
(i) $\lambda \leq 1$
(ii) $\lambda > \beta$ *iff*

$$\frac{D_C^R}{-d_R^R - (1+r)\,D_R^R}\left[(1-\beta)\,\zeta^R d_R^R + \zeta^D \left(d_R^D + \beta\,(1+r)\,D_R^D\right)\right] + \beta\zeta^D D_C^D < 0 \tag{3.12}$$

(iii) λ *increases in* D_C^R *and* $|d_R^R|$
(iv) λ *decreases in* d_R^D, D_R^D, D_C^D *and* $|D_R^R|$

As substitute types become cheaper in both periods, demand for the dirty backstop goes down in both periods. The strong green paradox arises if the damage from bringing forward exhaustible resource emissions $(1-\beta)\,\zeta^R d_R^R$ exceeds the benefits of reduced dirty backstop consumption in both periods. This is more likely when D_C^R is high, as a decrease in P^C then poses a larger threat to exhaustible resource demand in period 2. An increase in $|d_R^R|$ increases leakage by making it more attractive to shift exhaustible resource supply to period 1 (the reverse applies to $|D_R^R|$). Lastly, λ decreases in d_R^D, D_R^D and D_C^D, as high values of these parameters induce more substitution away from the dirty backstop.

By making stronger assumptions on the substitutability structure, we can obtain more powerful results about the occurrence of the green paradox and compare our findings with previous research. We analyze three special cases in which two of the energy types are perfect substitutes. When we calibrate Proposition 6.2 in section 3.5, we look at three scenarios that relate to these special cases.

3.4.1. Perfect substitutability between R and C

We are interested in this case as a reference point: the assumption that clean backstops are perfect substitutes for the exhaustible resource is common in green paradox models. It leads to the most powerful green paradox results in the literature. When the exhaustible resource and the green backstop are imperfect substitutes, exhaustible resource owners are ensured of future demand for their commodity and the green paradox may vanish (Gerlagh, 2011).

Corollary 3.1. *With perfect substitution between the exhaustible resource and the clean backstop*
(i) if $P^C > P^R$, a decrease in P^C has no effect
(ii) if $P^C = P^R$, then $\lambda > \beta$ if

$$(1-\beta)\,\zeta^R d_R^R + \zeta^D \left(d_R^D + \beta\,(1+r)\,D_R^D\right) < 0 \tag{3.13}$$

When P^C is sufficiently low, it fully determines exhaustible resource prices in both periods and the last term in (3.12) vanishes. In accordance with the literature, the condition for the strong green paradox is weaker than in the general case. Corollary 3.1 shows that if we take into consideration the availability of dirty backstops, the substitutability structure that is most conducive to the green paradox no longer suffices for its occurrence. Even when the exhaustible resource and the clean backstop are perfect substitutes, both near-term emissions and emission damages may go down as a result of lower clean backstop prices.

3.4.2. Perfect substitutability between D and C

Corollary 3.2. *With perfect substitution between the clean and the dirty backstop*
(i) if $P^C > P^D$, a decrease in P^C has no effect

(ii) if $P^C = P^D$, the strong green paradox does not occur

(iii) if $p^C > p^D$, $P^C < P^D$, then $\lambda > \beta$ iff

$$(1-\beta)\,\zeta^R d_R^R + \zeta^D d_R^D < 0 \tag{3.14}$$

(iv) if $p^C < p^D$ and $P^C < P^D$, then $\lambda = 1$

When the clean backstop is more expensive than the dirty backstop in both periods, the former is used in neither period and a small cost reduction has no effect. In the knife-edge case when the period 2 prices of the clean and the dirty backstop are equal, a reduction in the price of the clean backstop eliminates all demand for the dirty backstop, so there is no strong green paradox. When the clean backstop is already cheaper than the dirty backstop in the second period, further cost reductions only reduce dirty backstop use in period 1, at the cost of accelerated exhaustible-resource extraction. A green paradox then becomes more likely. When the clean backstop is cheaper than the dirty backstop in both periods, the latter is never used. The model reduces to a classic green paradox model and both the weak and the strong green paradox occur.

The analysis in Corollary 3.2 is complementary to Fischer and Salant (2010) and van der Ploeg and Withagen (2012b). Fischer and Salant (2010) analyze the effect of cheaper backstops in the presence of high- and low-cost oil. They find that moderate cost reductions for the backstop will cause some high-cost oil to remain unexploited and thus improve the environment. Beyond the point at which all high-cost oil remains in situ, further investments bring forward extraction of the low-cost oil and cause a 'renewed' green paradox. Van der Ploeg and Withagen (2012b) assume perfect substitutability between a clean and a dirty backstop and note that subsidizing renewables to the cost of the dirty backstop always reduces climate damages.

3.4.3. Perfect substitutability between R and D

Corollary 3.3. *With perfect substitution between the exhaustible resource and the dirty backstop*

(i) if $p^R < p^D$ and $P^R < P^D$, $\lambda = 1$

(ii) if $p^R < p^D$ and $P^R = P^D$, $\lambda = 0$

Table 3.2: Occurrence of the strong green paradox for cost reductions for the clean backstop

Substitutability between	Likelihood of strong GP	Related articles
R, C	+	Strand (2007); Hoel (2011); Gerlagh (2011)
D, C	-	van der Ploeg and Withagen (2012a) Fischer and Salant (2010)
R, D	-	

If the economy is in regime (ii), cost reductions benefit the environment by reducing the use of the dirty backstop in period 2, without affecting exhaustible resource extraction. When the clean backstop is sufficiently cheap, demand for the dirty backstop in period 2 goes to zero. The economy then moves into regime (i), in which additional investment only brings forward the extraction of the exhaustible resource and the green paradox returns.

Table 3.2 consolidates the discussion from the subsections.

3.5. Empirical calibration and special cases

In this section, we explore the magnitude of intertemporal carbon leakage by calibrating Propositions 6.4 and 6.2. When calibrating Proposition 6.2, we vary the substitutability between energy types and illustrate how the leakage rates change under different assumptions.

At the economy-wide level, conventional oil and coal are currently the biggest contributors to the climate change problem. We therefore approximate the effect of a future carbon tax on the time path of aggregate emissions by looking at its effect on oil and coal use. The substitution possibilities between oil and coal are not equally relevant in all sectors however. In electricity generation, coal primarily competes with conventional natural gas, the other scarce fossil fuel. We take a more detailed look at the electricity sector and calibrate the effect of a future tax on natural gas and coal emissions from electricity generation.

We observe current energy demand and prices, and the IEA forecasts future demand and prices. The upper panel of Table 3.3 presents an overview of these statistics. We

Table 3.3: Current and future energy demand and prices

	2009	2035
Aggregate oil demand	29165	36765
Aggregate coal demand	24096	38631
Oil price	60.4	135
Coal price	20.84	20.84
Natural gas demand for electricity generation	7337	12640
Coal demand for electricity generation	15683	26612
Natural gas price	40.14	78.12
Coal price	20.84	20.84

Quantities in mln boe, prices in 2009$ per boe. Emission intensities in t/boe: 0.3644 for oil; 0.5169 for coal; 0.3101 for natural gas. Values for 2035 are from the 'Current Policies' scenario of the World Energy Outlook 2010. A full list of definitions and sources is provided in Table 3.8.

use empirical estimates of interfuel own- and cross-price elasticities from previous work, though the values differ substantially across studies and may not adequately reflect long-term substitution possibilies (Stern, 2012).

In the first calibration exercise, we take oil as the exhaustible resource and coal as the dirty backstop. From the upper panel of Table 3.3, oil and coal demand are both expected to increase in 2035, though the relative increase is larger for coal. The oil price more than doubles during the next 26 years; we assume the coal price to remain constant. The literature on interfuel substitution typically distinguishes between coal and electricity as energy inputs. Since most coal is used for electricity generation, we take the elasticities for electricity as those for the dirty backstop.[41] We assume the elasticities to be equal across time periods.

We calculate the intertemporal leakage of a small global carbon tax in 2035. The first four columns in Table 3.4 contain the estimated economy-wide own- and cross-price elasticities for oil and electricity from five studies. Using the parameters from Table 3.3, for each set of estimates we determine the change in oil extraction and emissions and the intertemporal carbon leakage λ as a result of a small tax increase.

[41] We multiply the elasticity of oil demand with respect to electricity prices by the global share of coal-based electricity generation in total electricity generation, which was 0.47 in 2009 (IEA, 2011, p. 544).

Table 3.4: Estimates of economy-wide demand elasticities and intertemporal leakage predictions

Study	η_R^R	η_D^R	η_R^D	η_D^D	$\frac{\mathrm{d}d^R}{\mathrm{d}T}$	$\frac{\mathrm{d}e}{\mathrm{d}T}$	$\frac{\mathrm{d}E}{\mathrm{d}T}$	λ
Perkins (1994)	-0.25	0.07	0.11	-0.07	-18.93	-3.49	-15.56	-0.22
Cho et al. (2004)	-0.97	-0.01	-0.10	-0.79	48.67	19.93	-412.60	0.05
Ma et al. (2008)	-0.27	0.01	0.07	-0.68	9.07	2.27	-337.95	0.01
Serletis et al. (2009)[a]	-0.04	0.00	0.05	-0.07	0.63	-0.14	-30.88	-0.00
Serletis et al. (2010)	-0.12	0.04	0.07	-0.13	-10.23	-1.02	-50.46	-0.02

Equations for $\frac{\mathrm{d}d^R}{\mathrm{d}T}$, $\frac{\mathrm{d}e}{\mathrm{d}T}$ and $\frac{\mathrm{d}E}{\mathrm{d}T}$ are given by (3.30), (3.31) and (3.32), respectively. λ is defined in Definition 3.2. [a] Median estimates over all countries included in the study.

The intertemporal leakage rate is negative for three elasticity estimates, and the rates are small in absolute value. Except for Perkins' estimates, the effects of the future tax on second-period coal use through the own-price effect outweigh the substitution effects and the changes in the intertemporal pattern of oil extraction. For two sets of estimates in which η_D^R is high compared to $|\eta_R^R|$, the tax benefits oil exporters and increases second-period oil extraction, as in Persson et al. (2007): the tax-induced increase in future oil demand through substitution from coal to oil is larger than the decrease through higher own prices. The intertemporal leakage rate is negative for these two estimates: since oil is relatively cheap in the first period, the own-price effect of oil $\eta_R^R d^R/p^R$ is strong in the first period. When the oil price increases in the first period, the reduction in oil-related emissions exceeds the increase in coal-related emissions. Overall, the sum of period 1 and 2 emission reductions is almost linear in η_D^D, suggesting that the most important effect of carbon taxes is the direct reduction in coal use. The estimated emission reductions may be biased downwards, since we conservatively assumed that oil reserves are fully exhausted.

In the second calibration exercise, we look at the effects of a small future carbon tax on emissions in electricity generation, taking natural gas as the exhaustible resource. The lower panel of Table 3.3 shows current and future natural gas and coal demand for electricity generation. Demand for both energy types increases by 70% in 2035, though the relative price of natural gas goes up significantly.

Table 3.5 contains the leakage estimates using own- and cross-price elasticities for natural gas and coal in electricity generation. Natural gas demand is more elastic than

Table 3.5: Estimates of demand elasticities in electricity generation and intertemporal leakage predictions

Study	η_R^R	η_D^R	η_R^D	η_D^D	$\frac{\mathrm{d}d^R}{\mathrm{d}T}$	$\frac{\mathrm{d}e}{\mathrm{d}T}$	$\frac{\mathrm{d}E}{\mathrm{d}T}$	λ
Söderholm (2000)[a]	-0.82	0.03	0.04	-0.20	11.66	2.99	-70.75	0.04
Söderholm (2001)[a]	-0.38	0.03	0.09	-0.13	3.46	0.21	-42.26	0.00
Ko and Dahl (2001)	-1.46	1.54	0.28	-0.57	-150.45	-14.77	-78.47	-0.19
Serletis et al. (2010)	-0.14	0.14	0.06	-0.12	-13.85	2.91	-21.97	0.13

Equations for $\frac{\mathrm{d}d^R}{\mathrm{d}T}$, $\frac{\mathrm{d}e}{\mathrm{d}T}$ and $\frac{\mathrm{d}E}{\mathrm{d}T}$ are given by (3.30), (3.31) and (3.32), respectively. λ is defined in Definition 3.2. [a] Median estimates over all countries included in the study.

coal demand, as gas-fired power plants have higher marginal and lower fixed costs. Because natural gas is much cleaner than coal, a future tax increases future natural gas demand and delays natural gas extraction for two sets of elasticity estimates. When the cross-elasticities are small, as in Söderholm (2000) and Söderholm (2001), the tax accelerates natural gas extraction but the intertemporal leakage rate is low, as the direct decrease in second-period coal use dominates the effects on first-period emissions. The leakage rates have opposite signs in the bottom two rows depending on the $|\eta_R^R|/\eta_R^D$ ratio. For high values of this ratio (Ko and Dahl, 2001), higher first-period natural gas prices decrease first-period natural gas emissions by more than they increase emissions from coal. For low values of $|\eta_R^R|/\eta_R^D$ (Serletis et al., 2010), the converse applies and total first-period emissions increase.

Our calibration fixes the sum of current and future natural gas use for electricity generation. In reality, the change in cumulative natural gas use in electricity generation depends on the tax' effect on energy use in other sectors. Natural gas competes with oil, more than with coal, for heating and in industry (Stern, 2012). As oil is cleaner than coal, natural gas' comparative advantage vis-a-vis other fossil fuels as a result of a carbon tax is largest in electricity generation. By this reasoning, we may expect cumulative natural gas use in the electricity sector to increase. The resulting effect on emissions in the electricity sector depends on the elasticity of coal demand with respect to natural gas prices. A further caveat is that the difficulty to measure long-run substitution possibilities is especially relevant for electricity generation, as power plants have very long service lives. If we believe that the true own- and cross price elasticities are larger across the board, a future carbon tax will cause a larger reduction in cumulative emissions, but the leakage rate (which is a ratio of current and future emission reductions) may not be as responsive.

Finally, we calibrate Proposition 6.2. Energy types differ in their suitability for two main purposes: electricity generation and transport. Natural gas (R), coal (D) and wind and solar energy (C) more readily lend themselves for electricity generation, whereas oil (R), tar sands (D) and biofuels (C) are primarily used in the transportation sector. Energy types that are employed in the same submarket are closer substitutes than ones that are not. We look at three scenarios with different substitutability structures. The calibrations highlight the sensitivity of the intertemporal leakage rate to interfuel substitution possibilities, and illustrate how technology improvements for wind or solar energy and biofuels are likely to impact emissions from conventional oil, coal and unconventional oil - the three fossil fuels that present the biggest threat to the global climate.

The studies on interfuel substitution in the previous calibrations do not include renewable energy, so we follow the CGE literature on carbon leakage and assume a nested CES demand structure with two nests: electricity E and non-electricity N. We set the elasticity of substitution between nests at 1.5 and within nests at 5. In each scenario, we consider a relevant combination of energy types and group them by primary use. Appendix 3.A.2 contains a full description of the demand structure and parameter values per scenario.

3.5.1. Developing alternative fuels

In the first scenario, we study the effect of anticipated cost reductions for biofuels on emissions from oil (R) and coal (D). First-generation biofuels such as ethanol from sugarcane already compete with petroleum-based fuels, which currently dominate the transportation market. Second-generation fuels from biomass have the potential to be cheaper in the long run and exert less pressure on land and water supplies, but still face significant barriers to large-scale commercialization (Sims et al., 2010). By Proposition 6.2 and Corollary 3.1, the high substitutability between the clean backstop and the exhaustible resource and the low substitutability between the exhaustible resource and the dirty backstop are both associated with high leakage rates. This scenario is therefore likely to produce lavish estimates of the leakage rate.

Figure 3.1 shows the effect of reductions in the period 2 biofuel price on the first-period

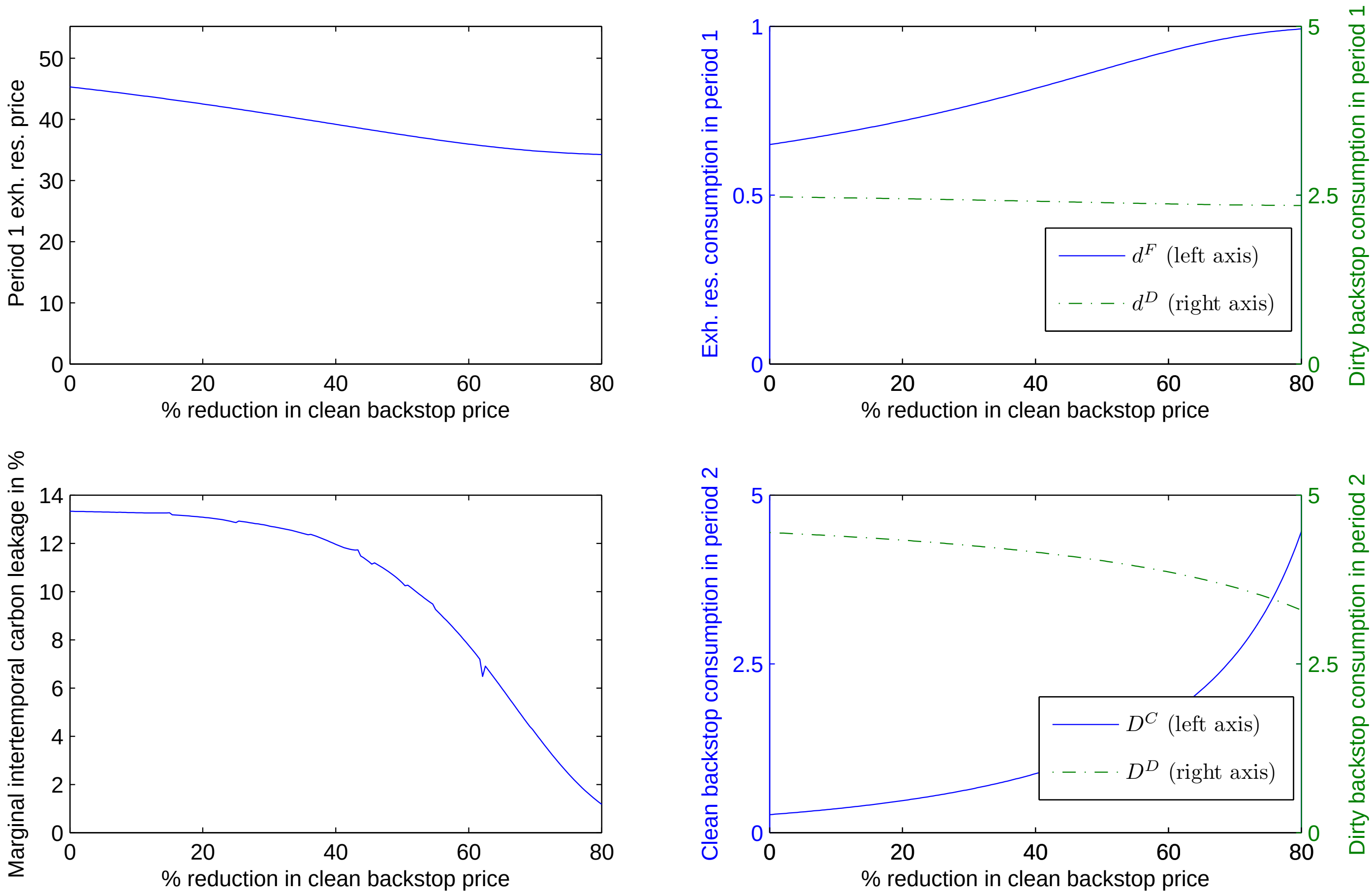

Figure 3.1: Effects of future cost reductions for biofuels (parametrization in Appendix 3.A.2).

oil price (top left), the marginal intertemporal leakage rate (bottom left), first-period oil and coal demand (top right) and second-period biofuel and coal demand (bottom right). The oil stock is normalized to one, so the top-right panel also indicates the fraction of the oil stock that is used in the first period. To facilitate comparisons of the change in first- and second-period coal demand, the right axes of the rightmost figures have the same scale. The horizontal axes indicate the relative second-period biofuel cost reduction as a fraction of the original second-period price, which is equal to the first-period price.

The share of the oil stock that is extracted in the first period increases almost linearly in the period 2 cost reduction for biofuels. For large cost reductions, the oil stock is exhausted almost completely in the first period. The associated drop in first-period oil prices is not large enough to substantially affect first-period coal consumption. When the cost reduction for biofuels is small, the marginal decrease in second-period coal use is also low, and the leakage rate reaches 13%. For larger cost reductions, it becomes comparatively more attractive to substitute biofuels for coal for non-transport purposes, as evidenced by the concave shape of the dashed curve in the bottom-right panel. The leakage rate drops to below 2% when the second-period cost reduction equals 80% of the initial price.

Of the renewable energy types, biofuels are the closest substitute to oil - the most emission intensive scarce fossil fuel. Investments in biofuel technologies are therefore more likely to lead to a green paradox than investments in wind or solar power. Still, the above calibration illustrates that biofuels' potential to reduce emissions from coal can reduce intertemporal leakage rates to levels significantly below the 100% that would obtain in models with one fossil fuel that is always fully exhausted. In a more comprehensive model that also includes unconventional oil, the leakage rate would be even lower, because unconventional oil is more abundant and more emission-intensive than conventional oil and a closer substitute for biofuels than coal is. On the other hand, we abstract from emissions from land use changes, which may lead to a downard bias in the leakage rate.

3.5.2. Renewable energy for electricity

In the second scenario, we consider oil (R) in the non-electricity nest and coal (D) and solar energy (C) in the electricity nest. The opportunities to employ renewable energy are highest in the electricity sector. Coal and renewable energy sources are main inputs for electricity generation, with worldwide market shares of 42% and 19% in 2008 respectively (IEA, 2010b). Investing in hydro-, wind- and solar power may reduce coal use without causing as strong an increase in short-term oil extraction.

Figure 3.2 depicts the effects of future cost reductions for solar energy. The oil price is unresponsive for small cost reductions, owing to the limited substitutability between oil and electricity. Only when solar energy becomes very cheap, it emerges as a credible substitute for oil and the oil price reacts more strongly. The pattern of intertemporal carbon leakage is similar. When the cost reduction is not too large, it mainly induces substitution from coal- to solar-based electricity in the second period. The decrease in second-period coal use dwarfs the change in oil extraction, giving rise to intertemporal leakage rates under 1%. The leakage rate goes up only when solar energy becomes very

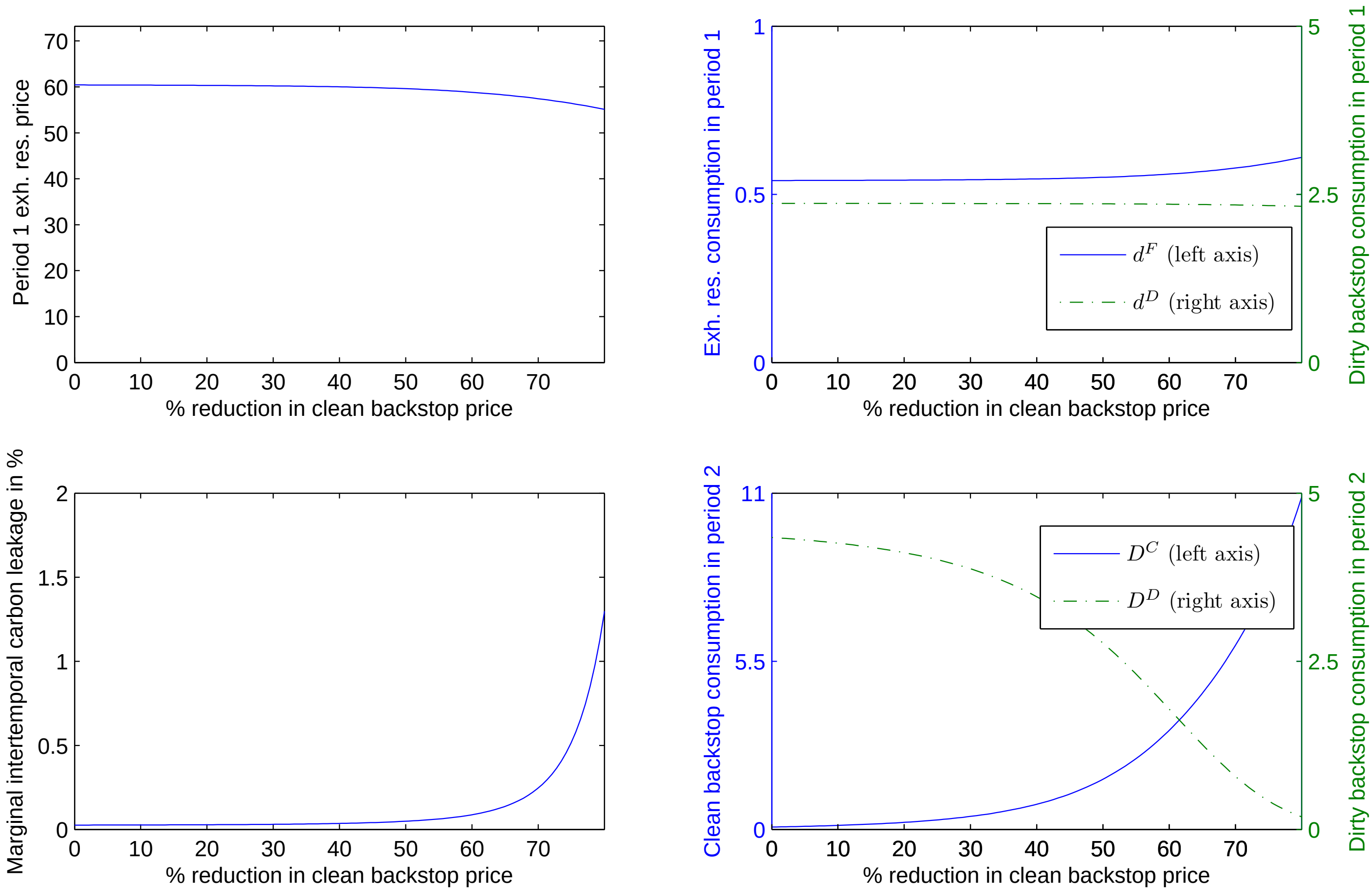

Figure 3.2: Effects of future cost reductions for solar-based electricity (parametrization in Appendix 3.A.2).

cheap in the future: coal is then hardly used anymore, and subsequent cost reductions increasingly serve to bring forward oil extraction. Extending the analysis to include natural gas may increase the leakage rate, though not by much as it is relatively emission extensive compared to coal.

This calibration complements the intuition behind Corollary 3.2, in which the clean and dirty backstop are perfect substitutes. From an environmental point of view, investment in renewable energy sources is primarily attractive insofar as it reduces the use of dirty backstops. When this goal has been achieved, intertemporal carbon leakage becomes a stronger concern. Compared to the first scenario, the high substitutability between the clean and the dirty backstop is favourable in terms of the leakage rate. The estimates from the second scenario suggest that improvements in renewable electricity technologies are especially suitable for reducing cumulative emissions, without bringing

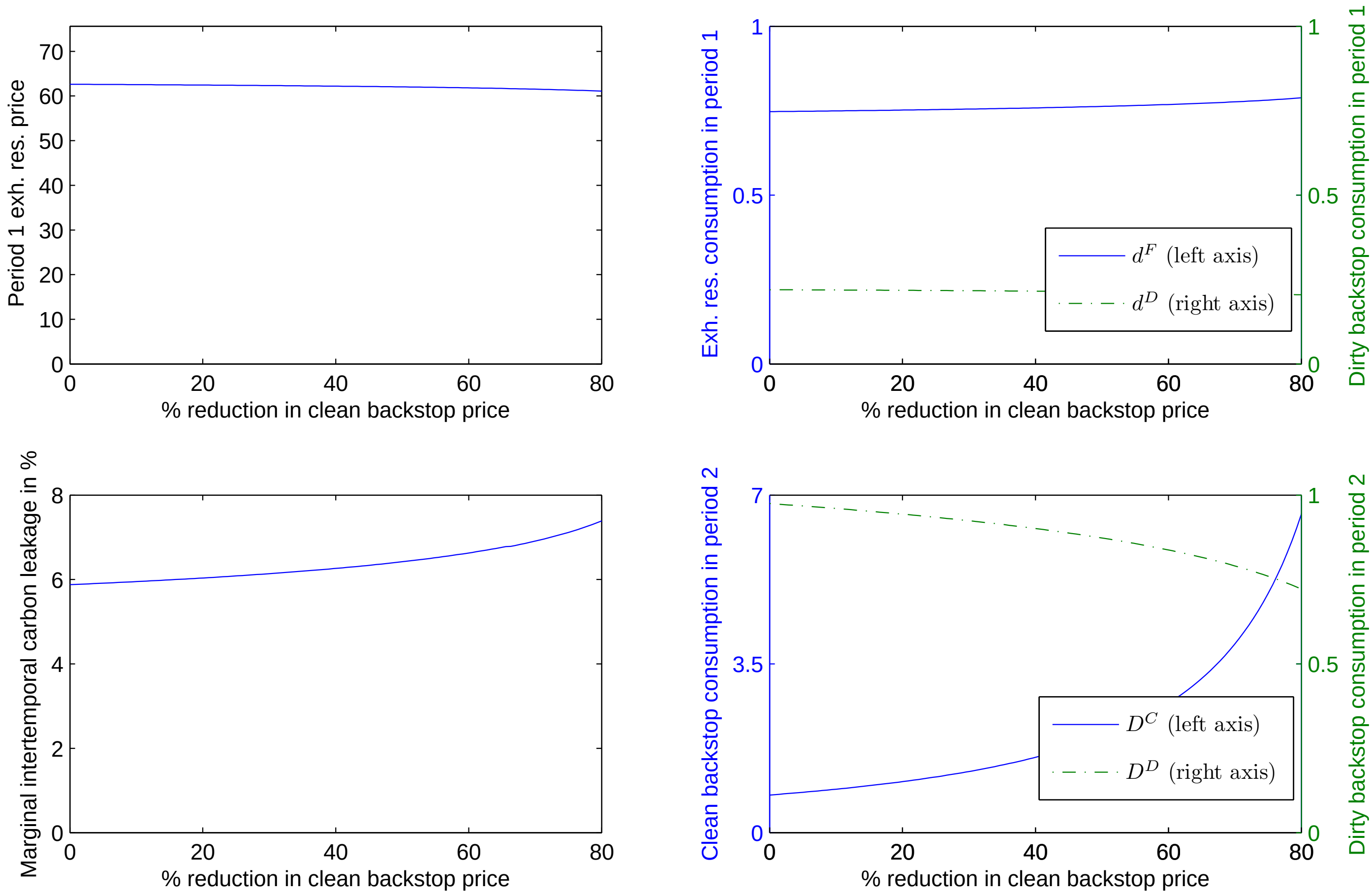

Figure 3.3: Effects of future cost reductions for solar-based electricity (parametrization in Appendix 3.A.2).

forward emissions from scarce fossil fuels.

3.5.3. Conventional and unconventional oil

In the last scenario, we illustrate how cost reductions for solar-based electricity (C) affect emissions from conventional (R) and unconventional (D) oil in the non-electricity nest.

Cost reductions for solar energy do not cause a big decrease in total second-period fossil fuel use, but the decrease is larger for unconventional oil: the slope of the dashed line in the bottom-right panel is larger than that of the solid line in the top-right panel. This is not surprising as unconventional oil is the dominant fuel in the second period, as stipulated by the Herfindahl rule. First-period emissions increase as first-period unconventional oil use is almost unaffected, but the leakage rates do not exceed 6-8%. These estimates accord with the prediction from Corollary 3.3 that reductions in future clean backstop prices primarily go at the expense of dirty backstop demand if the exhaustible resource and the dirty backstop are good substitutes.

3.6. Spatial carbon leakage

Climate policies may also have unintended side-effects in a spatial setting. When a group of countries reduces emissions unilaterally, pollution might move to other countries. This spatial carbon leakage occurs through two main channels (Felder and Rutherford, 1993). Firstly, dirty industries relocate to countries with laxer regulation. Secondly, a stringent environmental policy in environmentally conscious countries causes the world market price of fossil fuels to fall, increasing their use in lax countries. Estimated leakage rates range from a modest 2-5% to over 100%, the latter implying that unilateral carbon reduction policies increase global emissions (Burniaux and Oliveira Martins, 2012; Paltsev, 2001; Babiker, 2005).[42]

Our model can be used to analyze spatial rather than intertemporal carbon leakage by relabeling 'period 2' as a climate-conscious country, 'period 1' as a non-abating country and setting the interest rate to zero. The model does not explicitly incorporate industry relocation effects, but shows how unilateral carbon taxes affect world market fossil fuel prices and thereby emissions in non-abating countries. Paltsev (2001); Fischer and Fox (2009); Kuik and Hofkes (2010) argue that changes in energy prices are the most important determinant of carbon leakage.

Proposition 6.4 shows that spatial carbon leakage rates are below 100%. Reductions in dirty backstop use in abating countries are not offset by emission increases in non-abating countries. If unilateral climate policies decrease exhaustible resource prices, as policy makers sometimes fear, non-abating countries substitute away from dirty backstops, mitigating spatial carbon leakage. Leakage may even be negative if interfuel substitution possibilities differ between countries (as in Table 3.1). The assumption that total exhaustible-resource supply is fixed is less plausible than in the intertemporal model, but leads to conservative leakage estimates.

We calibrate a spatial version of the model to illustrate the magnitude of spatial

[42]Studies on international environmental agreements find a related effect. Because environmental standards are strategic substitutes, non-signatories will increase emissions (Barrett, 1994; Hoel, 1994).

Table 3.6: Aggregate demand and prices for spatial model

	ROW	EU
Aggregate oil demand	24959	4206
Aggregate coal demand	22143	1953
Oil price	60.40	68.00
Coal price	20.84	31.62

Quantities in mln boe, prices in 2009$ per boe. Emission intensities in t/boe: 0.3644 for oil; 0.5169 for coal. A full list of definitions and sources is provided in Table 3.8.

Table 3.7: Estimates of economy-wide demand elasticities and spatial leakage predictions

Study	η_R^R	η_D^R	η_R^D	η_D^D	$\frac{\mathrm{d}d^R}{\mathrm{d}T}$	$\frac{\mathrm{d}e}{\mathrm{d}T}$	$\frac{\mathrm{d}E}{\mathrm{d}T}$	λ
Perkins (1994)	-0.25	0.07	0.11	-0.07	0.52	0.09	-0.74	0.12
Cho et al. (2004)	-0.97	-0.01	-0.10	-0.79	19.98	8.25	-20.85	0.40
Ma et al. (2008)	-0.27	0.01	0.07	-0.68	4.87	1.18	-12.66	0.09
Serletis et al. (2009)[a]	-0.04	0.00	0.05	-0.07	0.57	-0.15	-1.06	-0.14
Serletis et al. (2010)	-0.12	0.04	0.07	-0.13	0.00	0.00	-1.69	0.00

Equations for $\frac{\mathrm{d}d^R}{\mathrm{d}T}$, $\frac{\mathrm{d}e}{\mathrm{d}T}$ and $\frac{\mathrm{d}E}{\mathrm{d}T}$ are given by (3.30), (3.31) and (3.32), respectively. λ is defined in Definition 3.2. [a] Median estimates over all countries included in the study.

leakage relative to intertemporal leakage. We disregard the time dimension and evaluate the leakage to the rest of the world (ROW) if the EU Emissions Trading Scheme (ETS) carbon price increases above its 2009 level of €15 per tonne. We restrict ourselves to the effects on aggregate oil and coal demand; we present a calibration for the electricity sector in the Appendix. Table 3.6 describes demand and prices in EU and ROW. The EU accounts for 14% of global oil demand and 8% of global coal demand. The EU ETS price adds about 50% to the price of coal.

Table 3.7 shows the effects of a carbon tax increase in the EU. The spatial leakage rates are mostly positive, unlike most intertemporal rates in Table 3.4. The EU consumes more oil compared to coal than the world at large, so the tax-induced reduction in EU oil demand is larger than the increase through substitution from coal to oil. A carbon tax increase in the EU therefore decreases world oil prices. Energy demand in ROW is similar to global energy demand. Like in the intertemporal calibration, a decrease in oil prices triggers an increase in emissions in ROW for four out of five estimates.

The spatial leakage rates tend to be larger than the intertemporal rates in absolute value, due to two reasons. Firstly, a tax increase has a modest effect on coal consumption in the EU, as coal demand in the EU is already low to begin with. This means that the denominator of the leakage rate - the total emission reduction in the EU - is low. Secondly, the future tax in the previous calibration is subject to a discount rate, whereas the unilateral tax is not. This exacerbates the reaction of oil suppliers in the three middle rows in Table 3.7, also increasing the absolute value of the leakage rate. For the estimates from Serletis et al. (2010), the direct and indirect effects of the EU tax on oil demand in the EU net out. Oil prices and hence ROW emissions remain unchanged, so the leakage rate is zero.

Our spatial leakage rates are similar to estimates from the computable general equilibrium literature on carbon leakage, albeit in the lower part of the spectrum (Di Maria and van der Werf, 2012). Comparing the intertemporal and spatial leakage rates, intertemporal leakage seems to be a small concern relative to spatial leakage.

Appendix 3.A.3 presents an extended model with two periods *and* two countries, in which one country implements a unilateral carbon tax in the second period. A future unilateral tax reduces second-period emissions in the adopting country. Future emissions in the non-abating country may increase or decrease, but any increase in emissions is lower than the reduction in the abating country. The effect of the future unilateral tax on first-period emissions is governed by a modified version of (3.9) which includes the first-period demand response in the non-abating country. The unweighted sum of current and future emissions in both countries always decreases as a result of a future unilateral tax. Emission damages can therefore only go up when total first-period emissions increase and the emission discount factor is sufficiently low.

3.7. Conclusion

We employ a general model to analyze intertemporal carbon leakage in the presence of an abundant dirty backstop such as coal or unconventional oil. The green paradox literature overstates the adverse consequences of imperfect climate policies by not taking into account their potential to reduce emissions from coal and unconventional oil. It is important to consider these fuels as they already account for 50% of energy-related emissions, and will become even more important in the future.

A carbon tax increases the price of oil, but the price of coal goes up even more. The effect of an anticipated carbon tax on future oil demand depends on the relative strength of a direct own-price and an indirect substitution effect. When improved technology (e.g. diesel from coal) makes coal a better substitute for oil in the future than it is today, intertemporal leakage may become negative. Anticipated carbon taxes cause significant substitution from coal to oil in the future and thereby induce oil owners to delay extraction. The reduction in present oil supply does not trigger a large increase in coal demand, as coal is a poor substitute for oil today. Future availability of cheap renewables lowers coal emissions directly (through substitution from coal to renewables) and indirectly (cost reductions for renewables decrease oil prices, reducing coal demand).

Calibrations of the model suggest that the effects of anticipated climate policies on present emissions are small compared to future emission reductions. Interestingly, some of the intertemporal leakage estimates are negative, implying that a future carbon tax reduces present emissions. Investments in renewable electricity technologies lead to very small leakage rates, because they are a good substitute for coal but have only a small impact on oil extraction.

The aim of climate policy is to decrease cumulative extraction, i.e. ensuring that some fossil fuels remain in the ground. The 'marginal resources' are not conventional oil and natural gas, which are so cheap to exploit that carbon taxes will only affect the distribution of rents and the timing of extraction. Climate policy should rather aim at reducing emissions from costly, emission-intensive and abundant resources such as coal and unconventional oil. The results from this paper imply that these efforts may be effective, even if it is not possible to instate all-encompassing carbon constraints.

3.A. Appendix

3.A.1. Proofs

Proof of Proposition 6.3

The weak green paradox occurs when (3.5) is positive. Using (3.6), the condition becomes

$$\frac{\partial e}{\partial p^R}\frac{\partial D^R}{\partial T} > 0 \tag{3.15}$$

The result follows by substituting (3.7) and (3.8).

Proof of Proposition 6.4

We first show that the tax increases the consumer price of the exhaustible resource by less than that of the dirty backstop.

Lemma 3.1. $\frac{dP^R}{dT} < \frac{dP^D}{dT}$

Proof. Assume not. Then, because own-price effects are stronger than cross-price effects (6.10), exhaustible resource demand in period 2 decreases. The tax increases the period 2 producer price $P^R - T\zeta^R$ as $\frac{\mathrm{d}P^D}{\mathrm{d}T} = \zeta^D$ and $\zeta^R < \zeta^D$. By the Hotelling condition (3.4), p^R increases and demand for the exhaustible resource goes down in period 1. This violates the requirement that the stock is fully exhausted. Q.E.D.

Lemma 6.1 and (6.10) entail $\frac{\mathrm{d}D^D}{\mathrm{d}T} < 0$, establishing (i). Totally differentiate with respect to T the stock constraint

$$\frac{\mathrm{d}d^R}{\mathrm{d}T} + \frac{\mathrm{d}D^R}{\mathrm{d}T} = d^R_R\frac{\mathrm{d}p^R}{\mathrm{d}T} + D^R_R\frac{\mathrm{d}P^R}{\mathrm{d}T} + D^R_D\zeta^Z = 0 \tag{3.16}$$

and the Hotelling condition (3.4)

$$\frac{\mathrm{d}p^R}{\mathrm{d}T} = \frac{1}{1+r}\left(\frac{\mathrm{d}P^R}{\mathrm{d}T} - \zeta^R\right) \tag{3.17}$$

Solving these two equations for $\frac{\mathrm{d}p^R}{\mathrm{d}T}$ and $\frac{\mathrm{d}P^R}{\mathrm{d}T}$ yields

$$\frac{\mathrm{d}p^R}{\mathrm{d}T} = -\frac{\zeta^R D_R^R + \zeta^D D_D^R}{d_R^R + (1+r)\, D_R^R} \tag{3.18}$$

$$\frac{\mathrm{d}P^R}{\mathrm{d}T} = \frac{\zeta^R d_R^R - (1+r)\,\zeta^D D_D^R}{d_R^R + (1+r)\, D_R^R} \tag{3.19}$$

The effect of T on E is

$$\begin{aligned}
\frac{\mathrm{d}E}{\mathrm{d}T} &= \zeta^D \frac{\mathrm{d}D^D}{\partial T} + \zeta^R \frac{\mathrm{d}D^R}{\mathrm{d}T} \\
&= \zeta^D \left(D_D^D \zeta^D + D_R^D \frac{\mathrm{d}P^R}{\mathrm{d}T} \right) + \zeta^R \left(D_R^R \frac{\mathrm{d}P^R}{\mathrm{d}T} + D_D^R \zeta^D \right) \\
&= \zeta^D \left(D_D^D \zeta^D + D_R^D \frac{\zeta^R d_R^R - (1+r)\,\zeta^D D_D^R}{d_R^R + (1+r)\, D_R^R} \right) + \zeta^R \left(D_R^R \frac{\zeta^R d_R^R - (1+r)\,\zeta^D D_D^R}{d_R^R + (1+r)\, D_R^R} + D_D^R \zeta^D \right) \\
&= \frac{1}{d_R^R + (1+r)\, D_R^R} \Big[(1+r) \left(\zeta^D\right)^2 \left(D_D^D D_R^R - D_R^D D_D^R \right) \\
&+ d_R^R \left(\left(\zeta^D\right)^2 D_D^D + \zeta^D \zeta^R \left(D_R^D + D_D^R \right) + \left(\zeta^R\right)^2 D_R^R \right) \Big]
\end{aligned} \tag{3.20}$$

The fraction outside the square brackets is negative. The first term inside the square brackets is positive by (6.10). The second term is also positive since $d_R^R < 0$ and

$$\begin{aligned}
&\left(\zeta^D\right)^2 D_D^D + \zeta^D \zeta^R D_D^R + \zeta^D \zeta^R D_R^D + \left(\zeta^R\right)^2 D_R^R < \\
&- \left(\zeta^D\right)^2 D_R^D + \zeta^D \zeta^R D_D^R + \zeta^D \zeta^R D_R^D - \left(\zeta^R\right)^2 D_D^R = \\
&- \left(\zeta^D D_R^D - \zeta^R D_D^R \right)^2 < 0
\end{aligned}$$

The first inequality follows from (6.10); the last equality from (3.A3). Therefore, $\frac{\mathrm{d}E}{\mathrm{d}T} < 0$, completing the proof of (ii). To prove (iii), we show that the sum of period 1 and 2 dirty backstop demand goes down.

$$\begin{aligned}
\frac{\mathrm{d}\left[d^D + D^D\right]}{\mathrm{d}T} &= d_R^D \frac{\mathrm{d}p^R}{\mathrm{d}T} + D_R^D \frac{\mathrm{d}P^R}{\mathrm{d}T} + \zeta^D D_D^D \\
&= d_R^D \frac{\mathrm{d}p^R}{\mathrm{d}T} + D_R^D \left((1+r) \frac{\mathrm{d}p^R}{\mathrm{d}T} + \zeta^R \right) + \zeta^D D_D^D \\
&= \frac{\mathrm{d}p^R}{\mathrm{d}T} \left(d_R^D + (1+r) D_R^D \right) + \zeta^R D_R^D + \zeta^D D_D^D \\
&= \underbrace{-\frac{d_R^D + (1+r) D_R^D}{d_R^R + (1+r) D_R^R}}_{<1} \underbrace{\left(\zeta^R D_R^R + \zeta^D D_D^R \right)}_{I} + \zeta^R D_R^D + \zeta^D D_D^D \\
&\overset{I>0}{<} \zeta^R \left(D_R^R + D_R^D \right) + \zeta^D \left(D_D^R + D_D^D \right) < 0
\end{aligned} \tag{3.21}$$

The second equality in (3.21) follows by substituting (3.17); the fourth by substituting (3.18). The fraction in (3.21) is smaller than one by (6.10). The second inequality holds by (6.10) and (3.A3). When $I < 0$, $\frac{\mathrm{d}\left[d^D + D^D\right]}{\mathrm{d}T}$ is negative by (6.10), completing the proof of (iii). Lastly, calculate the effect on emission damages

$$\begin{aligned}
\frac{\mathrm{d}\Sigma}{\mathrm{d}T} &= (1-\beta) \zeta^R \frac{\mathrm{d}d^R}{\mathrm{d}T} + \zeta^D \left(d_R^D \frac{\mathrm{d}p^R}{\mathrm{d}T} + \beta \left(D_R^D \frac{\mathrm{d}P^R}{\mathrm{d}T} + \zeta^D D_D^D \right) \right) \\
&= (1-\beta) \zeta^R \frac{\mathrm{d}d^R}{\mathrm{d}T} + \zeta^D \left(\frac{\mathrm{d}p^R}{\mathrm{d}T} \left(d_R^D + \beta (1+r) D_R^D \right) + \beta \left(\zeta^R D_R^D + \zeta^D D_D^D \right) \right) \\
&= \frac{\mathrm{d}p^R}{\mathrm{d}T} \left[(1-\beta) \zeta^R d_R^R + \zeta^D \left(d_R^D + \beta (1+r) D_R^D \right) \right] + \beta \zeta^D \left(\zeta^R D_R^D + \zeta^D D_D^D \right) \\
&= -\frac{\zeta^R D_R^R + \zeta^D D_D^R}{d_R^R + (1+r) D_R^R} \left[(1-\beta) \zeta^R d_R^R + \zeta^D \left(d_R^D + \beta (1+r) D_R^D \right) \right] + \\
&\quad \beta \zeta^D \left(\zeta^R D_R^D + \zeta^D D_D^D \right)
\end{aligned} \tag{3.22}$$

The second equality in (3.22) follows from (3.17); the fourth by substituting (3.18).

Proof of Proposition 6.1

The change in first-period emissions is

$$\frac{\mathrm{d}e}{\mathrm{d}P^C} = \zeta^R d_R^R \frac{\mathrm{d}p^R}{\mathrm{d}P^C} + \zeta^D d_R^D \frac{\mathrm{d}p^R}{\mathrm{d}P^C} = \frac{\mathrm{d}p^R}{\mathrm{d}P^C} \left(\zeta^R d_R^R + \zeta^D d_R^D \right) \tag{3.23}$$

Analogously to the proof of Proposition 6.4, we can back out $\frac{\mathrm{d}p^R}{\mathrm{d}P^C}$. Totally differentiate the stock constraint and (3.4) with respect to P^C

$$d_R^R \frac{\mathrm{d}p^R}{\mathrm{d}P^C} + D_R^R \frac{\mathrm{d}P^R}{\mathrm{d}P^C} + D_C^R = 0 \tag{3.24}$$

$$\frac{\mathrm{d}p^R}{\mathrm{d}P^C} = \frac{1}{1+r} \frac{\mathrm{d}P^R}{\mathrm{d}P^C} \tag{3.25}$$

Solving for $\frac{\mathrm{d}p^R}{\mathrm{d}P^C}$, we obtain

$$\frac{\mathrm{d}p^R}{\mathrm{d}P^C} = -\frac{D_C^R}{d_R^R + (1+r)\, D_R^R} \tag{3.26}$$

The weak green paradox occurs when $\frac{\mathrm{d}e}{\mathrm{d}P^C} < 0$. As $\frac{\mathrm{d}p^R}{\mathrm{d}P^C} > 0$, $\frac{\mathrm{d}e}{\mathrm{d}P^C} < 0$ iff $\zeta^R d_R^R + \zeta^D d_R^D < 0$.

Proof of Proposition 6.2

We established that $\frac{\mathrm{d}p^R}{\mathrm{d}P^C} > 0$, so by (3.4), exhaustible resource prices in both periods are increasing in P^C. Then d^D, D^D, $e + E$ and E are increasing in P^C. It follows that $\lambda = -\frac{\mathrm{d}e}{\mathrm{d}P^C} / \frac{\mathrm{d}E}{\mathrm{d}P^C} \leq 1$, proving (i). The change in emission damages is

$$\begin{aligned}
\frac{\mathrm{d}\Sigma}{\mathrm{d}P^C} &= (1-\beta)\, \zeta^R \frac{\mathrm{d}d^R}{\mathrm{d}P^C} + \zeta^D \left(d_R^D \frac{\mathrm{d}p^R}{\mathrm{d}P^C} + \beta \left(D_R^D \frac{\mathrm{d}P^R}{\mathrm{d}P^C} + D_C^D \right) \right) \\
&= (1-\beta)\, \zeta^R \frac{\mathrm{d}d^R}{\mathrm{d}P^C} + \zeta^D \left(d_R^D \frac{\mathrm{d}p^R}{\mathrm{d}P^C} + \beta \left(D_R^D (1+r) \frac{\mathrm{d}p^R}{\mathrm{d}P^C} + D_C^D \right) \right) \\
&= \frac{\mathrm{d}p^R}{\mathrm{d}P^C} \left[(1-\beta)\, \zeta^R d_R^R + \zeta^D \left(d_R^D + \beta (1+r)\, D_R^D \right) \right] + \beta \zeta^D D_C^D \\
&= \frac{D_C^R}{-d_R^R - (1+r)\, D_R^R} \left[(1-\beta)\, \zeta^R d_R^R + \zeta^D \left(d_R^D + \beta (1+r)\, D_R^D \right) \right] + \beta \zeta^D D_C^D
\end{aligned} \tag{3.27}$$

The second equality in (3.27) follows from (3.25); the last from (3.26). The leakage rate is

$$\begin{aligned}
\lambda &= -\frac{\frac{\mathrm{d}p^R}{\mathrm{d}P^C}\left(\zeta^R d_R^R + \zeta^D d_R^D\right)}{\zeta^R \frac{\mathrm{d}D^R}{\mathrm{d}P^C} + \zeta^D \left(D_R^D \frac{\mathrm{d}P^R}{\mathrm{d}P^C} + D_C^D\right)} \\
&= \frac{\frac{D_C^R}{d_R^R + (1+r) D_R^R}\left(\zeta^R d_R^R + \zeta^D d_R^D\right)}{\zeta^R \frac{d_R^R D_C^R}{d_R^R + (1+r) D_R^R} + \zeta^D \left(\frac{D_R^D}{D_R^R}\left(\frac{d_R^R D_C^R}{d_R^R + (1+r) D_R^R} - D_C^R\right) + D_C^D\right)} \\
&= \frac{D_C^R \left(\zeta^R d_R^R + \zeta^D d_R^D\right)}{\zeta^R d_R^R D_C^R + \zeta^D \left(-D_R^D (1+r) D_C^R + D_C^D \left(d_R^R + (1+r) D_R^R\right)\right)}
\end{aligned} \tag{3.28}$$

The second equality follows from $\frac{\mathrm{d}D^R}{\mathrm{d}P^C} = -\frac{\mathrm{d}d^R}{\mathrm{d}P^C}$ and (3.26). By taking derivatives of (3.28), we obtain (iii) and (iv).

Proof of Corollary 3.1

We omit the proof of (i). For (ii), note that the price of the clean backstop fully determines exhaustible resource prices

$$\begin{aligned}
&\lim_{\left(D_R^R, D_R^C\right) \to (-\infty, +\infty)} \frac{\mathrm{d}P^R}{\mathrm{d}P^C} = \\
&\lim_{\left(D_R^R, D_R^C\right) \to (-\infty, +\infty)} \left(-\frac{(1+r) D_C^R}{d_R^R + (1+r) D_R^R}\right) = 1
\end{aligned} \tag{3.29}$$

The D_C^D term in (3.27) is superfluous because of (3.29) and since dirty backstop users are indifferent between substituting to the exhaustible resource and to the clean backstop. $\frac{\mathrm{d}\Sigma}{\mathrm{d}P^C}$ then has the same sign as the term in square brackets in (3.27).

Proof of Corollary 3.2

We omit the proof of (i). For (ii), D^D is infinitely elastic with respect to P^C at $P^C = P^D$. From (3.28), we see that $\lim_{D_C^D \to \infty} \lambda = 0$. For (iii), $D_C^D = D_D^R = 0$ when $P^C < P^D$. It then follows that $\frac{\mathrm{d}\Sigma}{\mathrm{d}P^C}$ has the same sign as $(1-\beta)\,\zeta^R d_R^R + \zeta^D d_R^D$. For (iv), $d_R^D = D_C^D = D_D^R = 0$ when $p^C < p^D$ and $P^C < p^D$. We then have $\frac{\mathrm{d}\Sigma}{\mathrm{d}P^C} < 0$.

Proof of Corollary 3.3

The proof of (i) is analogous to the proof of (iv) in Corollary 3.2. For (ii), we see in (3.28) that $\lim_{(D_R^R, D_R^D)\to(-\infty,+\infty)} \lambda = 0$.

3.A.2. Calibrations

In Table 3.4, the expressions for $\frac{\mathrm{d}d^R}{\mathrm{d}T}$, $\frac{\mathrm{d}e}{\mathrm{d}T}$ and $\frac{\mathrm{d}E}{\partial T}$ are given by $d_R^R \frac{\mathrm{d}p^R}{\mathrm{d}T}$, (3.15) and (3.20) respectively.[43] Substituting elasticities for the partial derivatives of the demand function and using the assumption that the elasticities are equal across periods, we get

$$\frac{\mathrm{d}d^R}{\mathrm{d}T} = \frac{\zeta^R \eta_R^R \dfrac{D^R}{P^R} + \zeta^D \eta_D^R \dfrac{D^R}{P^D}}{-1 - (1+r)\dfrac{D^R}{P^R}\dfrac{p^R}{d^R}} \tag{3.30}$$

$$\frac{\mathrm{d}e}{\mathrm{d}T} = \frac{\mathrm{d}d^R}{\mathrm{d}T}\left(\zeta^R + \zeta^D \frac{\eta_R^D}{\eta_R^R}\frac{d^D}{d^R}\right) \tag{3.31}$$

$$\frac{\mathrm{d}E}{\mathrm{d}T} = \left(\zeta^D\right)^2 \eta_D^D \frac{D^D}{P^D} + \zeta^D \eta_R^D \frac{D^D}{P^R}\left(\frac{-\frac{\mathrm{d}d^R}{\mathrm{d}T}}{\eta_R^R \frac{D^R}{P^R}} + \zeta^D \frac{\eta_D^R}{-\eta_R^R}\frac{P^R}{P^D}\right) - \zeta^R \frac{\mathrm{d}d^R}{\mathrm{d}T} \tag{3.32}$$

Table 3.8 lists the definitions and sources of the variables used for the calibrations in Tables 3.4 and 3.7. In the calibrations for spatial leakage, P^R and P^D are inclusive of a €15 per tonne carbon tax.

Table 3.9 contains information on natural gas and coal use for electricity generation in the EU and in the rest of the world (ROW). The EU relies more on natural gas for its electricity generation vis-a-vis ROW. Because natural gas is only 60% as emission intensive as coal, the relative price difference for the two resources is much smaller in the EU.

Table 3.10 presents the spatial leakage calibration for the electricity sector. As in the intertemporal calibration in Table 3.5, a unilateral tax increase raises natural gas demand for electricity generation in the EU for the two bottom elasticity estimates. Because the EU consumes more natural gas and less coal relative to the world at large, the direct own-price effect of the tax on natural gas consumption is stronger than in the intertemporal calibration, and the substitution effect from coal to gas is weaker. For Ko and Dahl (2001) and Serletis et al. (2010), the reaction of natural gas suppliers is therefore weaker than in the intertemporal calibration, bringing the leakage rates closer

[43] The derivation of these equations does not depend on assumption (3.A3).

Table 3.8: Data definitions and sources for emission tax calibrations

Variable	Definition	Source
	Table 3.4	
d^R, d^D	Total primary oil and coal demand in 2009, mln boe	IEA (2011, p. 544)
D^R, D^D	Total primary oil and coal demand in 2035, mln boe	IEA (2010b, p. 619)
p^R	IEA crude oil import price in 2009, \$ per barrel	IEA (2010b, p. 71)
P^R	IEA crude oil import price in 2035, \$ per barrel	IEA (2010b, p. 71)
p^D, P^D	Average EU steam coal import costs in 2009, \$ per boe[a]	IEA (2010a, p. III.44)
	Table 3.5	
d^R, d^D	Natural gas and coal demand for power generation in 2009, mln boe	IEA (2011, p. 544)
D^R, D^D	Natural gas and coal demand for power generation in 2035, mln boe	IEA (2010b, p. 619)
p^R	European natural gas import import price in 2009, \$ per boe	IEA (2010b, p. 71)
P^R	European natural gas import import price in 2035, \$ per boe	IEA (2010b, p. 71)
p^D, P^D	Average EU steam coal import costs in 2009, \$ per boe	IEA (2010a, p. III.44)
	Table 3.7	
d^R, d^D	EU total primary oil and coal demand in 2009, mln boe	IEA (2011, p. 564)
D^R, D^D	ROW total primary oil and coal demand in 2009, mln boe	IEA (2011, p. 544, p. 564)
p^R	IEA crude oil import price in 2009, \$ per boe	IEA (2010b, p. 71)
p^D	Average EU steam coal import costs in 2009, \$ per boe	IEA (2010a, p. III.44)
P^i	$p^i + €15^b \zeta^i,\ i \in \{R, D\}$	
	Table 3.10	
d^R, d^D	EU natural gas and coal demand for power generation in 2009, mln boe	IEA (2011, p. 544)
D^R, D^D	ROW natural gas and coal demand for power generation in 2035, mln boe	IEA (2011, p. 544, p. 564)
p^R	European natural gas import import price in 2009, \$ per boe	IEA (2010b, p. 71)
p^D	Average EU steam coal import costs in 2009, \$ per boe	IEA (2010a, p. III.44)
P^i	$p^i + €15^a \zeta^i,\ i \in \{R, D\}$	
ζ^i	$\dfrac{\text{Total CO2 emissions from type } i \text{ in 2009, t}}{\text{Total primary energy demand for type } i \text{ in 2009, boe}}$	IEA (2011, p. 544, p. 546)

All energy demand is converted from Mtoe in source data; 1 toe = 7.315 boe. Energy prices are converted from \$ per MBtu or \$ per tce in source data; 1 MBtu = 0.1843 boe; 1 tce = 4.787 boe. All 2035 values are from the 'Current Policies' scenario in IEA (2010b).
[a] Exchange rate: \$1 = €0.719

Table 3.9: Demand for electricity generation and prices for spatial model

	ROW	EU
Natural gas demand for electricity generation	6320	1017
Coal demand for electricity generation	14147	1536
Natural gas price	40.14	46.63
Coal price	20.84	31.62

Quantities in mln boe, prices in 2009$ per boe. Emission intensities in t/boe: 0.3101 for natural gas; 0.5169 for coal. A full list of definitions and sources is provided in Table 3.8.

Table 3.10: Estimates of demand elasticities in electricity generation and spatial leakage predictions

Study	η_R^R	η_D^R	η_R^D	η_D^D	$\frac{\mathrm{d}d^R}{\mathrm{d}T}$	$\frac{\mathrm{d}e}{\mathrm{d}T}$	$\frac{\mathrm{d}E}{\mathrm{d}T}$	λ
Söderholm (2000)[a]	-0.82	0.01	0.04	-0.20	4.67	1.18	-3.86	0.31
Söderholm (2001)[a]	-0.38	0.01	0.09	-0.13	2.02	0.10	-1.92	0.05
Ko and Dahl (2001)	-1.46	0.72	0.28	-0.57	-1.87	-0.16	-5.30	-0.03
Serletis et al. (2010)	-0.14	0.07	0.06	-0.12	-0.16	0.04	-1.20	0.03

Equations for $\frac{\mathrm{d}d^R}{\mathrm{d}T}$, $\frac{\mathrm{d}e}{\mathrm{d}T}$ and $\frac{\mathrm{d}E}{\mathrm{d}T}$ are given by (3.30), (3.31) and (3.32), respectively. λ is defined in Definition 3.2. [a] Median estimates over all countries included in the study.

to zero. The spatial leakage rate for the Söderholm (2000) estimates is higher than the intertemporal rate in Table 3.5 because the tax' direct effect on coal consumption in the EU is low, owing to the low initial level of coal consumption in the EU.

For the calibrations in sections 3.5.1, 3.5.2 and 3.5.3, we employ the following model. Let X denote composite energy and indicate nests by $k \in \{N, E\}$. The elasticity of substitution between and within nests is σ_X and σ_k, respectively. The value share of nest k in composite energy demand is α_k^X; the share of type i in nest k is α_i^k. Denote available income by y. Then

$$d^i = \left(\frac{y}{p^X}\right) \sum_{k \in \{N,E\}} \alpha_k^X \left(\frac{p^X}{p^k}\right)^{\sigma_X} \alpha_i^k \left(\frac{p^i}{p^k}\right)^{\sigma_k} \tag{3.33}$$

$$p^X = \left(\sum_{k \in \{N,E\}} \alpha_k^X \left(p^K\right)^{\frac{1-\sigma_X}{\sigma_X}}\right)^{\frac{\sigma_X}{1-\sigma_X}}, \; p^k = \left(\sum_{i \in \{R,D,C\}} \alpha_i^k \left(p^i\right)^{\frac{1-\sigma_k}{\sigma_k}}\right)^{\frac{\sigma_k}{1-\sigma_k}} \tag{3.34}$$

where $\sum_{k \in \{N,E\}} \alpha_k^X = \sum_{i \in \{R,D,C\}} \alpha_i^k = 1$. Demand in the second period is described by a similar system. Exhaustible resource prices p^R and P^R are endogenously determined by (3.4) and $d^R + D^R = S$. The parameter values are listed in Table 6.1. As (3.34) is homogeneous of degree zero in S, y and Y, we normalize S to one. Income is chosen such that $Y = (1 + r)\, y$ and $p^R = 60.4$ when $P^C = p^C$ in the second scenario. The interest rate equals 2% per annum compounded over 26 years. The coal price and emission intensities for coal and conventional oil are equal to those in Table 3.8. Unconventional oil is 20% more emission-intensive than conventional oil. We set the unconventional oil price and the initial biofuel price to 80.

3.A.3. Spatial and intertemporal leakage

In this section, we extend the model to include both a spatial and an intertemporal dimension, to analyze the effect on emissions when one region implements a unilateral carbon tax T_2 in the second period. We denote variables corresponding to the adopting country with capitals, and indicate time by subscripts. Then $D_{D,1}^R$ is the derivative of exhaustible resource demand with respect to the dirty backstop price in the adopting country in the first period. Exhaustible resource owners must now be indifferent between

Table 3.11: Parametrization of calibrations in sections 3.5.1, 3.5.2 and 3.5.3

Variable	Interpretation			
ζ^R	Emission intensity of exhaustible resource		0.3644	
r	Interest rate		0.6734	
S	Exhaustible resource stock		1	
y	First-period income		84.03	
Y	Second-period income		140.59	
σ_X	Elasticity of substitution between nests		1.5	
			Section	
		3.5.1	3.5.2	3.5.3
σ_N	Elasticity of substitution within non-electricity nest	5		5
σ_E	Elasticity of substitution within electricity nest		5	
α_N^X	Value share of non-electricity in aggregate energy	0.5	0.5	0.2
α_R^N	Value share of exhaustible resource in non-electricity nest	0.5	1	0.5
α_D^N	Value share of dirty backstop in non-electricity nest	0	0	0.5
α_C^N	Value share of clean backstop in non-electricity nest	0.5	0	0
α_R^E	Value share of exhaustible resource in electricity nest	0	0	0
α_D^E	Value share of dirty backstop in electricity nest	1	0.5	0
α_C^E	Value share of clean backstop in electricity nest	0	0.5	1
p^D	Dirty backstop price	20.84	20.84	80
p^C	Initial clean backstop price	80	46.25	46.25
ζ^D	Emission intensity of dirty backstop	0.5169	0.5169	0.4373

selling in either region in either period:

$$p_1^R = P_1^R = \frac{1}{1+r} p_2^R = \frac{1}{1+r}\left(P_2^R - T_2 \zeta^R\right) \tag{3.35}$$

Similar to (3.8), the unilateral second-period tax increases exhaustible resource prices in the first period and in the non-adopting country if the tax increases exhaustible resource demand at constant producer prices

$$\frac{\mathrm{d}p_1^R}{\mathrm{d}T_2}, \frac{\mathrm{d}P_1^R}{\mathrm{d}T_2}, \frac{\mathrm{d}p_2^R}{\mathrm{d}T_2} \gtreqless 0 \Leftrightarrow \zeta^R D_{R,2}^R + \zeta^D D_{R,2}^D \gtreqless 0 \tag{3.36}$$

Proposition 3.5. *Following a unilateral carbon tax increase in the second period*
(i) D_2^D *decreases*
(ii) under (3.A3), E_2 *decreases*
(iii) under (3.A3), $e_1 + E_1 + e_2 + E_2$ *decreases*
(iv) $e_1 + E_1$ *increases when* $\left(\zeta^R D_{R,2}^R + \zeta^D D_{D,2}^R\right)\left(\zeta^R\left(d_{R,1}^R + D_{R,1}^R\right) + \zeta^D\left(d_{R,1}^D + D_{R,1}^D\right)\right) >$
0
(v) under (3.A3), $e_2 + E_2$ *decreases*

Proof. The proof of (i) and (ii) is analogous to Proposition 6.4. Totally differentiating (3.35) with respect to T_2 yields

$$\frac{\mathrm{d}p_1^R}{\mathrm{d}T_2} = \frac{\mathrm{d}P_1^R}{\mathrm{d}T_2} = \frac{1}{1+r}\frac{\mathrm{d}p_2^R}{\mathrm{d}T_2} = \frac{1}{1+r}\left(\frac{\mathrm{d}P_2^R}{\mathrm{d}T_2} - \zeta^R\right) \tag{3.37}$$

From $d_1^R + D_1^R + d_2^R + D_2^R = S$, we get

$$d_{R,1}^R \frac{\mathrm{d}p_1^R}{\mathrm{d}T_2} + D_{R,1}^R \frac{\mathrm{d}P_1^R}{\mathrm{d}T_2} + d_{R,2}^R \frac{\mathrm{d}p_2^R}{\mathrm{d}T_2} + D_{R,2}^R \frac{\mathrm{d}P_2^R}{\mathrm{d}T_2} + D_{D,2}^R \zeta^D = 0 \tag{3.38}$$

Solve (3.37) and (3.38) in $\frac{\mathrm{d}p_1^R}{\mathrm{d}T_2}$, $\frac{\mathrm{d}P_1^R}{\mathrm{d}T_2}$, $\frac{\mathrm{d}p_2^R}{\mathrm{d}T_2}$ and $\frac{\mathrm{d}P_2^R}{\mathrm{d}T_2}$

$$\frac{\mathrm{d}p_1^R}{\mathrm{d}T_2} = -\frac{\zeta^R D_{R,2}^R + \zeta^D D_{D,2}^R}{\Gamma} \tag{3.39a}$$

$$\frac{\mathrm{d}P_1^R}{\mathrm{d}T_2} = -\frac{\zeta^R D_{R,2}^R + \zeta^D D_{D,2}^R}{\Gamma} \tag{3.39b}$$

$$\frac{\mathrm{d}p_2^R}{\mathrm{d}T_2} = -\frac{(1+r)\left(\zeta^R D_{R,2}^R + \zeta^D D_{D,2}^R\right)}{\Gamma} \tag{3.39c}$$

$$\frac{\mathrm{d}P_2^R}{\mathrm{d}T_2} = \frac{\left(d_{R,1}^R + D_{R,1}^R + (1+r)\,d_{R,2}^R\right)\zeta^R - \zeta^D(1+r)\,D_{D,2}^R}{\Gamma} \tag{3.39d}$$

where $\Gamma \equiv d_{R,1}^R + D_{R,1}^R + (1+r)\,d_{R,2}^R + (1+r)\,D_{R,2}^R$. Analogous to Proposition 6.4, we prove (iii) by showing that total dirty backstop use goes down.

$$\begin{aligned}
\frac{\mathrm{d}\left[d_1^D + D_1^D + d_2^D + D_2^D\right]}{\mathrm{d}T_2} =& d_{R,1}^D\frac{\mathrm{d}p_1^R}{\mathrm{d}T_2} + D_{R,1}^D\frac{\mathrm{d}P_1^R}{\mathrm{d}T_2} + d_{R,2}^D\frac{\mathrm{d}p_2^R}{\mathrm{d}T_2} + D_{R,2}^D\frac{\mathrm{d}P_2^R}{\mathrm{d}T_2} + \zeta^D D_{D,2}^D \\
=& -d_{R,1}^D\frac{\zeta^R D_{R,2}^R + \zeta^D D_{D,2}^R}{\Gamma} - D_{R,1}^D\frac{\zeta^R D_{R,2}^R + \zeta^D D_{D,2}^R}{\Gamma} \\
& -(1+r)\,d_{R,2}^D\frac{\zeta^R D_{R,2}^R + \zeta^D D_{D,2}^R}{\Gamma} \\
& + D_{R,2}^D\frac{\left(d_{R,1}^R + D_{R,1}^R + (1+r)\,d_{R,2}^R\right)\zeta^R - \zeta^D(1+r)\,D_{D,2}^R}{\Gamma} + \zeta^D D_{D,2}^D \\
=& \frac{1}{\Gamma}\left[\zeta^D D_{D,2}^D\left(d_{R,1}^R + D_{R,1}^R + (1+r)\,d_{R,2}^R + (1+r)\,D_{R,2}^R\right)\right. \\
& -\zeta^D D_{D,2}^R\left(d_{R,1}^D + D_{R,1}^D + (1+r)\,d_{R,2}^D + (1+r)\,D_{R,2}^D\right) \\
& -\zeta^R D_{R,2}^R\left(d_{R,1}^D + D_{R,1}^D + (1+r)\,d_{R,2}^D\right) \\
& \left.+\zeta^R D_{R,2}^D\left(d_{R,1}^R + D_{R,1}^R + (1+r)\,d_{R,2}^R\right)\right] \\
& < \frac{1}{\Gamma}\left[\underbrace{\zeta^D D_{D,2}^D\left(d_{R,1}^R + D_{R,1}^R + (1+r)\,d_{R,2}^R\right)}_{I}\right. \\
& \underbrace{-\zeta^D D_{D,2}^R\left(d_{R,1}^D + D_{R,1}^D + (1+r)\,d_{R,2}^D\right)}_{II} \\
& \underbrace{-\zeta^R D_{R,2}^R\left(d_{R,1}^D + D_{R,1}^D + (1+r)\,d_{R,2}^D\right)}_{III} \\
& \left.\underbrace{+\zeta^R D_{R,2}^D\left(d_{R,1}^R + D_{R,1}^R + (1+r)\,d_{R,2}^R\right)}_{IV}\right]
\end{aligned} \tag{3.40}$$

The last inequality follows from (6.10). In (3.40), $I,\ III > 0$ and $II,\ IV < 0$. When the tax decreases oil demand in the adopting country in the second period, that is if

$\zeta^R D^R_{R,2} + \zeta^D D^R_{D,2} \leq 0$, we have $II + III \geq 0$. From $I + IV > 0$ (by (6.10)) it then follows that $\frac{\mathrm{d}\left[d_1^D + D_1^D + d_2^D + D_2^D\right]}{\mathrm{d}T_2} < 0$. Conversely, when $\zeta^R D^R_{R,2} + \zeta^D D^R_{D,2} > 0$, (6.10) gives

$$\begin{aligned}
\frac{\mathrm{d}\left[d_1^D + D_1^D + d_2^D + D_2^D\right]}{\mathrm{d}T_2} <& \frac{1}{\Gamma}\left[\zeta^D D^D_{D,2}\left(d^R_{R,1} + D^R_{R,1} + (1+r)\, d^R_{R,2}\right)\right.\\
&+\zeta^D D^R_{D,2}\left(d^R_{R,1} + D^R_{R,1} + (1+r)\, d^R_{R,2}\right)\\
&+\zeta^R D^R_{R,2}\left(d^R_{R,1} + D^R_{R,1} + (1+r)\, d^R_{R,2}\right)\\
&\left.+\zeta^R D^D_{R,2}\left(d^R_{R,1} + D^R_{R,1} + (1+r)\, d^R_{R,2}\right)\right]\\
=& \frac{\overbrace{d^R_{R,1} + D^R_{R,1} + (1+r)\, D^R_{R,2}}^{>0}}{\Gamma}\\
&\underbrace{\left(\zeta^D D^D_{D,2} + \zeta^D D^R_{D,2} + \zeta^R D^R_{R,2} + \zeta^R D^D_{R,2}\right)}_{<0} < 0
\end{aligned}$$

The underbraced term is negative by (6.10) and (3.A3), completing the proof of (iii). The effect of the tax on first-period emissions is

$$\begin{aligned}
\frac{\mathrm{d}\left[e_1 + E_1\right]}{\mathrm{d}T_2} =& \frac{\mathrm{d}p_1^R}{\mathrm{d}T_2}\left(\zeta^R d^R_{R,1} + \zeta^D d^D_{R,1}\right) + \frac{\mathrm{d}P_1^R}{\mathrm{d}T_2}\left(\zeta^R D^R_{R,1} + \zeta^D D^D_{R,1}\right)\\
=& -\frac{\left(\zeta^R D^R_{R,2} + \zeta^D D^R_{D,2}\right)\left(\zeta^R d^R_{R,1} + \zeta^D d^D_{R,1}\right)}{\Gamma} - \frac{\left(\zeta^R D^R_{R,2} + \zeta^D D^R_{D,2}\right)\left(\zeta^R D^R_{R,1} + \zeta^D D^D_{R,1}\right)}{\Gamma}\\
=& -\frac{\left(\zeta^R D^R_{R,2} + \zeta^D D^R_{D,2}\right)\left(\zeta^R\left(d^R_{R,1} + D^R_{R,1}\right) + \zeta^D\left(d^D_{R,1} + D^D_{R,1}\right)\right)}{\Gamma}
\end{aligned} \tag{3.41}$$

As $\Gamma < 0$, first-period emissions increase when the numerator in (3.41) is positive, establishing (iv). The effect of the tax on second-period emissions is

$$\begin{aligned}
\frac{\mathrm{d}\left[e_2+E_2\right]}{\mathrm{d}T_2} =& \frac{\mathrm{d}e_2}{\mathrm{d}T_2}+\frac{\mathrm{d}E_2}{\mathrm{d}T_2} \\
=& \zeta^D\frac{\mathrm{d}d_2^D}{\mathrm{d}T_2}+\zeta^R\frac{\mathrm{d}d_2^R}{\mathrm{d}T_2}+\frac{\mathrm{d}E_2}{\mathrm{d}T_2} \\
=& \zeta^D d_{R,2}^D\frac{\mathrm{d}p_2^R}{\mathrm{d}T_2}+\zeta^R\frac{\mathrm{d}d_2^R}{\mathrm{d}T_2}+\frac{\mathrm{d}E_2}{\mathrm{d}T_2} \\
=& \frac{\mathrm{d}p_2^R}{\mathrm{d}T_2}\left(\zeta^D d_{R,2}^D+\zeta^R d_{R,2}^R\right)+\frac{\mathrm{d}E_2}{\mathrm{d}T_2} \\
=& \frac{\mathrm{d}p_2^R}{\mathrm{d}T_2}\left(\zeta^D d_{R,2}^D+\zeta^R d_{R,2}^R\right)+\zeta^D\left(D_{D,2}^D\zeta^D+D_{R,2}^D\frac{\mathrm{d}P^R}{\mathrm{d}T}\right)+\zeta^R\left(D_{R,2}^R\frac{\mathrm{d}P^R}{\mathrm{d}T}+D_{D,2}^R\zeta^D\right) \\
=& -\frac{(1+r)\left(\zeta^R D_{R,2}^R+\zeta^D D_{D,2}^R\right)\left(\zeta^R d_{R,2}^R+\zeta^D d_{R,2}^D\right)}{\Gamma}+\left(\zeta^D\right)^2 D_{D,2}^D+\zeta^D\zeta^R D_{D,2}^R \\
&+\frac{\left(\zeta^D D_{R,2}^D+\zeta^R D_{R,2}^R\right)\left(\left(d_{R,1}^R+D_{R,1}^R+(1+r)\,d_{R,2}^R\right)\zeta^R-\zeta^D(1+r)\,D_{D,2}^R\right)}{\Gamma} \\
=& \frac{1}{\Gamma}\Bigg[\zeta^D(1+r)\Bigg[\overbrace{\zeta^D D_{D,2}^D D_{R,2}^R-\zeta^D D_{R,2}^D D_{D,2}^R}^{>0}-\overbrace{\zeta^D d_{R,2}^D D_{D,2}^R}^{I}-\overbrace{\zeta^R d_{R,2}^D D_{R,2}^R}^{II}+\overbrace{\zeta^D D_{D,2}^D d_{R,2}^R}^{III} \\
&+\overbrace{\zeta^R D_{R,2}^D d_{R,2}^R}^{IV}\Bigg]+\underbrace{\left(d_{R,1}^R+D_{R,1}^R\right)}_{<0}\underbrace{\left[\left(\zeta^D\right)^2 D_{D,2}^D+\zeta^D\zeta^R D_{D,2}^R+\zeta^D\zeta^R D_{R,2}^D+\left(\zeta^R\right)^2 D_{R,2}^R\right]}_{V}\Bigg]
\end{aligned} \tag{3.42}$$

The fifth equation follows by substituting the analog of (3.20) and the sixth from (3.39). When $d_{R,2}^R=-d_{R,2}^D$

$$-I-II+III+IV=d_{R,2}^D\left(-\zeta^D D_{D,2}^R-\zeta^R D_{R,2}^R-\zeta^D D_{D,2}^D-\zeta^R D_{R,2}^D\right)>0$$

The inequality follows from (3.A3). Because $III+IV>0$, by induction we also have $-I-II+III+IV>0$ when $d_{R,2}^R<-d_{R,2}^D$. For V, we have

$$\begin{aligned}
&\left(\zeta^D\right)^2 D_{D,2}^D+\zeta^D\zeta^R D_{D,2}^R+\zeta^D\zeta^R D_{R,2}^D+\left(\zeta^R\right)^2 D_{R,2}^R< \\
&-\left(\zeta^D\right)^2 D_{R,2}^D+\zeta^D\zeta^R D_{D,2}^R+\zeta^D\zeta^R D_{R,2}^D-\left(\zeta^R\right)^2 D_{D,2}^R= \\
&-\left(\zeta^D D_{R,2}^D-\zeta^R D_{D,2}^R\right)^2<0
\end{aligned}$$

The first inequality follows from (6.10); the last equality from (3.A3). As $\Gamma<0$, $\frac{\mathrm{d}[e_2+E_2]}{\mathrm{d}T_2}<0$. <u>Q.E.D.</u>

Chapter 4

Strategic resource extraction and substitute development[44]

4.1. Introduction

The oil market has flavours of a bilateral monopoly. The largest exporters have united themselves in OPEC, a cartel that controls more than 75% of proven reserves[45] and actively manages supply. Importing countries coordinate on energy policy and energy security issues through various international organizations such as the IEA, OECD and EU, and cooperate in the development of renewable alternatives. Consuming countries are vulnerable to monopoly power because of their heavy dependency on oil, but also have the means to end this dependency through developing a backstop: a substitute that can replace oil as the dominant energy source.

To prolong consumers' dependence on their resource, oil exporters have an incentive to prevent prices from becoming too high. Indeed, one of OPEC's aims is to 'secure an efficient, economic and regular supply of petroleum to consumers'. Conversely, importing countries realize that investment in alternative energy sources, such as shale gas or renewable energy, is not only a remedy to the physical scarcity of oil, but also affects exporters' supply decisions. The resulting strategic interaction is the subject of this paper: we ask how an exhaustible resource seller can adjust the supply path to preserve its monopoly, and how a buyer can optimally use the ability to develop a substitute.

A key feature of exhaustible resource markets is that expectations about future demand affect current supply. A binding promise by the buyer in the initial period about

[44] This chapter also circulates as Michielsen (2013c).

[45] Known reserves that with reasonable certainty can be recovered in the future under existing economic and operating conditions. Source: BP Statistical Review of World Energy 2012.

the arrival time of the substitute (Dasgupta et al., 1983; Gallini et al., 1983) will therefore not only change market conditions when the substitute comes on line, but also affect the supply path in all preceding periods (Karp and Newbery, 1993). As time passes, the effect on supply in early periods becomes sunk, so the buyer faces a different trade-off than in the initial period. As a result, the buyer's optimal open-loop strategy is not time-consistent (Olsen, 1986, 1993): the buyer has an incentive to commit to late development in order to depress current prices, but would like to renege on his promise when the resource becomes scarce (see section 4.3). Sellers who are sufficiently rational to calculate dynamic equilibria are unlikely to be naive enough to believe announcements that are not credible. In order to present a realistic model of their supply decision, it is important to exclude non-credible promises by the buyer. The contribution of this paper is to derive the closed-loop solution to the investment and supply game.

We use a simple model and a highly stylized representation of the innovation process. In each period, a monopolistic seller makes a supply offer to a buyer. After observing the seller's offer, the buyer decides whether to pay a fixed investment cost and develop a perfect substitute. Upon investment, the substitute can immediately be produced at constant marginal cost and competes with the resource. We abstract from uncertainty, capacity constraints, R&D externalities and imperfect cartelization in order to focus on the strategic aspects of the resource supply and innovation decisions.

In equilibrium, the seller induces the buyer to delay the adoption of the substitute until the resource is exhausted. When the buyer cannot commit to future actions, his only means to influence current supply is to invest immediately. Doing so adversely affects the buyer in three ways. Firstly, the seller immediately reduces supply following investment. Secondly, the buyer loses the ability to derive surplus from the resource through saving the interest on the investment cost. Thirdly, a larger share of the remaining resource stock is sold at the buyer's reservation price. From the seller's perspective, the buyer's ability to develop a substitute is equivalent to an already available substitute with a higher marginal cost. Like in Hoel (1978), the seller limit-prices at the buyer's reservation price, which is higher than the marginal cost of the substitute because of the fixed investment cost. The buyer's indifference condition for investment only becomes binding when the remaining resource stock is sufficiently small.

Our model offers an explanation for the slow progress in renewable energy. Developing a substitute does not help industrialized countries to capture a larger share of the oil rent. So as long as oil prices do not become prohibitively high, there is little economic rationale for substantive efforts.

Four recent papers consider similar questions. Liski and Montero (2011) study the optimal demand schedule of an exhaustible resource monopsonist that has access to a substitute. Their results are similar to ours to the extent that the buyer would like to commit to postpone the switch to the substitute to depress current prices, and the buyer's share of the resource surplus increases in the cost of the substitute when the buyer cannot commit. An important difference is that the perfectly competitive sellers in Liski and Montero (2011) only compete with the substitute for an infinitesimally short period. Because the exhaustion of the resource always coincides with the switch to the substitute, postponing the switch decreases current resource supply. The buyer obtains lower prices because committing to late invention increases his monopsony power. In our framework, the monopolistic seller may compete with the substitute for a nondegenerate period when the buyer can commit the invention time (Olsen, 1993). Postponing the arrival of the substitute increases cumulative resource supply before invention. Current prices go down because current supply goes up, rather than because of an increase in monopsony power as in Liski and Montero (2011). As a result, initial resource use is higher under commitment than under discretion in our paper, but lower in Liski and Montero (2011).[46]

In Harris and Vickers (1995), the arrival of the substitute is a stochastic event and depends on the buyer's R&D effort. As the buyer's effort increases when the stock dwindles, the seller has an incentive to slow down depletion in closed-loop equilibrium. Jaakkola (2012) looks at a game with gradual substitute development and convex per-period investment costs. This generalization prohibits an analytical characterization of the closed-loop equilibrium. In Gerlagh and Liski (2011), the substitute is not available immediately following investment, but after an exogenous transition period. By investing, the buyer can force the seller to sell the remaining stock during the transition period. The value of this investment option decreases as the resource is depleted. To compensate the buyer for this decrease, supply is increasing over time during an interval

[46] A formal proof for the general case is beyond the scope of this paper. We provide an example with iso-elastic demand in Appendix 4.A.3.

before investment. Although Gerlagh and Liski look for a Markov-perfect equilibrium, the time-to-build delay acts as a commitment device.

4.2. Model

Consider a model with a buyer and a seller of an exhaustible resource. We adopt the notation of Gerlagh and Liski (2011). The buyer derives flow surplus $u(q_t)$ from consuming q_t units of the resource. Let $u(.)$ be increasing and continuously differentiable. The surplus function is associated with a consumer utility function $\widetilde{u}\,(.)$ satisfying $u\,(q_t) = \widetilde{u}\,(q_t) - \widetilde{u}'\,(q_t)\,q_t$ and inverse demand function $\psi(q_t) = \widetilde{u}'\,(q_t)$. A monopolistic seller is endowed with a finite stock s_0 that is costless to extract. The seller aims to maximize the discounted stream of instantaneous profits $\pi\,(q_t) = \widetilde{u}'\,(q_t)\,q_t$. Utility and profits are discounted at rate r.

The buyer can instantaneously develop a perfect substitute for the resource. By paying an upfront investment cost I, he gains the ability to produce the substitute at constant marginal cost c. The buyer is then guaranteed a flow surplus

$$\bar{u} \equiv u\left(\psi^{-1}\,(c)\right)$$

There are $i = 1, .., N$ periods of time of length ε. Time is continuous but agents only act at the beginning of each period, i.e. at time points $t_i = \varepsilon(i-1)$. They commit their actions for the entire period. We distinguish between a pre-investment phase A and a post-investment phase B. In phase A, the buyer has a binary decision variable $k_{t_i} \in \{0, 1\}$, where $k_{t_i} = 1$ indicates that the buyer develops the substitute. In phase A, each period has three stages:

(1) the seller offers to supply a quantity q_{t_i}
(2) the buyer chooses $k_{t_i} \in \{0, 1\}$
(3) when $k_{t_i} = 0$, the market clears at price $p_{t_i} = \psi(q_{t_i})$, or when $k_{t_i} = 1$, the seller's offer is rejected. The seller makes a revised offer q'_{t_i} and the market clears at price $p_{t_i} = \min\left(\psi(q'_{t_i}), c\right)$. At the end of the period, the economy moves to regime B.

In phase B, the seller is the only mover and chooses quantity q_{t_i}. The market clears at price $p_{t_i} = \min(\psi(q_{t_i}), c)$, because the buyer cannot prevent the seller from competing with the substitute.

Our timing reflects the mutual dependence of buyers and sellers in natural resource markets, and oil exporters' concerns about security of supply: the seller may be able to dissuade the buyer from investing by making a sufficiently generous offer. Olsen (1993) uses the same setup as our paper, but the buyer moves before the seller in each period (stages (1) and (2) are reversed). This timing forces the buyer to invest very early, as he cannot credibly punish the seller for setting a high price.[47] We will characterize the limit of the equilibrium when ε goes to zero. For convenience, the remainder of the analysis is therefore presented in continuous time with an infinite horizon.[48][49]

4.3. The time-consistency problem

Olsen (1993) sketches why equilibria in which the buyer can commit the investment time are not consistent. According to the Hotelling rule, a monopolistic seller equates discounted marginal revenue in all periods. When the arrival time of the substitute T' is fixed, additional supply in any period after T' does not cause inframarginal losses because the resource is sold at the substitute price. As marginal revenue is relatively high after the arrival of the substitute, the seller would like to sell part of his stock in the post-investment phase. This is socially inefficient and reduces the buyer's welfare, because the seller directly displaces substitute supply using scarce resource units. When the buyer commits to a late T', the seller's post-investment revenues go down in present

[47] Olsen does not analyze large initial stocks and restricts himself to the limit-pricing phase. For large stocks, the buyer cannot credibly prevent the seller from treating his problem as a free-terminal-time problem with a scrap value, the scrap value being the discounted profits from the limit-pricing phase. The buyer thus invests before the remaining stock reaches the region that Olsen considers. Because the resource is exhausted in finite time, the folk theorem cannot sustain Pareto dominant outcomes. The crucial difference between the timing in this paper and in Olsen (1993) is that the buyer moves in between the seller's proposal and the market clearance. Reversing stages (2) and (3) would yield the same equilibrium as in Olsen (1993): after the seller's offer in the initial period, the residual game is equivalent to Olsen (1993).

[48] The buyer's ability to develop the substitute guarantees that the strategic interaction ends in finite time. When N is sufficiently large, changes in N do not affect the equilibrium. Moreover, when the number of remaining periods is large, the equilibrium strategies do not depend on the number of remaining periods.

[49] In general, it is not clear whether the limit of the equilibrium is an equilibrium of the limit game; see Simon and Stinchcombe (1989) for an example.

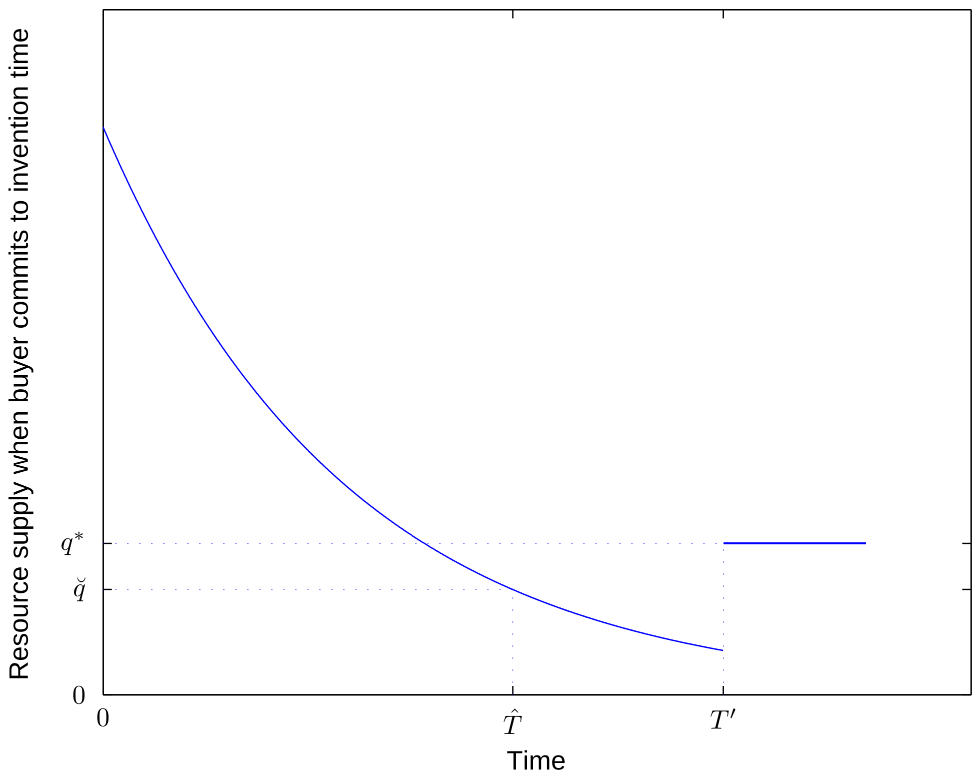

Figure 4.1: The buyer's time-consistency problem

terms because of discounting.

A late T' results in higher cumulative extraction in the pre-investment phase and increased supply in early periods, but is costly for the buyer in later periods: when the resource becomes scarce, the buyer cannot dispose over the substitute as early as he would like. After having enjoyed the benefits of high supplies and low prices early on, the buyer has an incentive to avoid the costs of resource scarcity in subsequent periods by developing the substitute earlier than announced. We illustrate this incentive in Figure 4.1. After $\hat{T}$, supply drops below the buyer's reservation quantity $\breve{q}$, which is defined such that $u(\breve{q})$ equals the utility $u(q^*)$ of consuming the long-run substitute supply q^* minus the interest on the investment cost rI. After $\hat{T}$, the buyer prefers to invest immediately instead of at the announced time T'.[50]

[50] A formal treatment of the buyer's commitment optimum is beyond the scope of this paper, and is available in Dasgupta et al. (1983); Olsen (1986).

4.4. Post-investment phase

When making decisions during phase A, agents take into account their payoffs in phase B. We therefore first turn to phase B. When the period length ε is sufficiently small, we can approximate the seller's problem in phase B by its continuous-time analogue:

$$\begin{aligned}
&\max_{q_t} \int_0^\infty \min\left(\psi\left(q_t\right), c\right) q_t e^{-rt} dt \\
&\text{s.t. } \dot{s} = -q_t,\ q_t \geq 0,\ s_t \geq 0,\ s_0 \text{ given}
\end{aligned} \tag{4.1}$$

The Hamiltonian for this problem and the necessary optimality conditions are

$$\begin{aligned}
&H\left(q, s, \lambda\right) = \min\left(\psi\left(q_t\right), c\right) q_t - \lambda q \\
&\pi'\left(q_t\right) - \lambda_t \leq 0,\ q \geq \psi^{-1}\left(c\right) \text{ c.s.} \\
&\dot{\lambda}_t = r\lambda_t,\ \lim_{t\to\infty} \lambda_t s_t = 0
\end{aligned} \tag{4.2}$$

where λ is the scarcity value of the resource and

$$\pi'\left(q_t\right) = \begin{cases} \psi\left(q_t\right) - q_t\psi'\left(q_t\right) & \text{for } q_t \geq D\left(c\right) \\ c & \text{for } q_t < D\left(c\right) \end{cases}$$

Because a perfect substitute is available at cost c, marginal revenue $\pi'\left(q_t\right)$ is equal to c for the first $q^* \equiv \psi^{-1}\left(c\right)$ units and discontinuous at q^*, as the seller starts incurring losses on inframarginal units. If the seller practices a limit-pricing strategy, supplying q^* in each period until exhaustion, the stock is exhausted at time s/q^*. For small values of the remaining stock, the marginal revenue $\pi'\left(q_t\right)|_{q_t \geq q^*}$ of supplying a unit in excess of q^* is always lower than the present value of c at time s/q^*. The seller then supplies q^* in each period until exhaustion time T. When the remaining stock is large, limit pricing takes so long to exhaust the stock that $\pi'\left(q^*\right)$ is higher than ce^{-rs_t/q^*}. It is then worthwhile to supply $q_t > q^*$ in early periods.

Proposition 4.1 (post-investment phase, Hoel (1978))**.** *Let*

$$s^* \equiv -\frac{q^*}{r}\ln\left(\frac{\pi'(q^*)}{c}\right)$$
$$\Delta \equiv \frac{s^*}{q^*}$$

When the remaining stock $s < s^*$

$$q_t = q^* \ \forall\ t \in (0, T)\,,\ T = \frac{s}{q^*}$$

When $s > s^*$

$$q_t = \pi'^{-1}\left(ce^{-r(T-t)}\right)\ \forall\ t \in [0, T-\Delta)$$
$$q_t = q^* \ \forall\ t \in [T-\Delta, T] \tag{4.3}$$
$$s.t. \int_0^{T-\Delta} \pi'^{-1}\left(ce^{-r(T-t)}\right) dt + \Delta q^* = s \tag{4.4}$$

The above system defines the seller's optimal post-investment strategy $\phi^I(s;c)$*.*

Figure 6.1 illustrates the seller's supply path during the post-investment phase. In equilibrium, the seller is indifferent between supplying an extra unit in any period before $T-\Delta$ and marginally extending the limit-pricing period. The present value of marginal revenue $\pi'(q_t)\,e^{-rt}$ equals ce^{-rT} between $[0, T-\Delta]$ and at T.

4.5. Closed-loop equilibrium

It will be useful to define the value function for the seller $V(s)$ and buyer $U(s)$ as a function of the remaining stock during phase A. Denote the buyer's and seller's value at the time of investment by $U^I(s)$ and $V^I(s)$, respectively. Let $k = \kappa(s, q)$ be the buyer's investment strategy and $q = \phi(s)$ the seller's supply strategy. The value for the seller is the payoff from choosing the optimal q for an interval of length ε, given the strategic response of the buyer

$$V(s) = \max_{\{q\}}\{[\varepsilon\pi(q) + e^{-\varepsilon r}V(s-\varepsilon q)](1-\kappa(s,q)) + V^I(s)\kappa(s,q)\}. \tag{4.5}$$

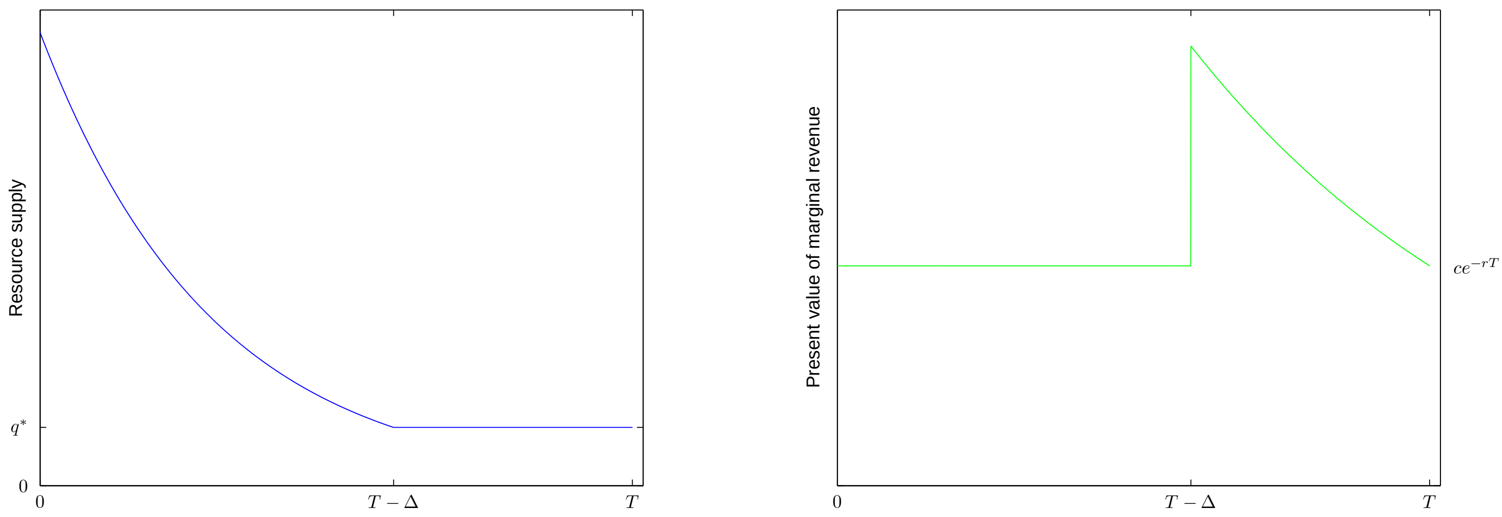

Figure 4.2: The seller's optimal supply (left) and the present value of marginal revenue (right) during the post-investment phase

We define the value for the buyer analogously

$$U(s) = \max_{k\in\{0,1\}} \{[\varepsilon u(\phi(s)) + e^{-\varepsilon r}U(s-\varepsilon\phi(s))](1-k) + U^I(s)k\}. \tag{4.6}$$

Because the buyer can guarantee himself the long-run surplus $\bar{u} - rI$ by investing, the buyer's strategy maximizes his value of the resource $W(s)$, i.e. the utility the resource provides on top of his long-run level

$$W(s) = U(s) - \left(\frac{\bar{u}}{r} - I\right)$$

The post-investment values $U^I(s)$ and $V^I(s)$ are determined by the dynamics described in Proposition 4.1.

$$W^I(s) = \int_0^{T-\Delta} [u(q_t) - \bar{u}]\, e^{-rt} dt \tag{4.7}$$

We look for a pair of equilibrium strategies such that $\phi(s) = \text{argmax}_{\{q\}} V(s)$ and $\kappa(s,q) = \text{argmax}_k U(s)$ for all s. Proposition 4.2 states the main result.

Proposition 4.2. *There exists a subgame-perfect equilibrium in pure strategies. The* *buyer's equilibrium strategy is*

$$\kappa(s,q) = 1 \ \forall \ q : u(q) < \bar{u} - rI \tag{4.8a}$$

$$\kappa(s,q) = 0 \ o.w. \tag{4.8b}$$

The seller's equilibrium strategy is $\phi(s) = \phi^I\left(s; \psi\left(u^{-1}\left(\bar{u} - rI\right)\right)\right)$, with $\phi^I(s;c)$ as defined in Proposition 4.1.[51]

Although the motivation differs in the pre- and post-investment phase, the seller always supplies at least the buyer's reservation quantity. In the post-investment phase, the seller always supplies $q \geq q^*$ because there is no loss on inframarginal units for smaller q. In the pre-investment phase, the seller's best response to the buyer's equilibrium strategy is to supply at least $q \geq \breve{q}$, because he would trigger investment for smaller q. From the seller's perspective, the threat of the substitute is therefore equivalent to an already available substitute with a higher marginal cost.

The buyer does not invest before exhaustion because his surplus from the resource W decreases after investment. Adopting the substitute raises the buyer's reservation utility from $\bar{u} - rI$ to $\bar{u}$, but not his actual utility as the buyer pays interest rI on the sunk investment cost. The increase in the buyer's reservation utility has three effects, as we show in Figure 4.3. Firstly, the buyer's surplus from the resource ($u(q)$ minus the reservation utility) is lower for any given supply path: $u(q) - u(\breve{q}) > u(q) - \bar{u}$. Secondly, the seller decreases supply in early periods when the buyer invests: $\phi(s) > \phi^I(s)$ when $s > s^*$ (Hoel, 1978, 1983). The maximum price for the resource decreases after investment; to compensate for this, the seller increases the price in early periods.

[51] Because the buyer conditions his strategy on q, there can also exist a large number of other equilibria. To see this, consider a very small c and a non-negligible I. The seller's surplus after investment $V^I(s)$ is very small, so the buyer can, for all levels of the remaining stock, threaten to invest unless the seller offers the buyer's first-best quantity for that s. When faced with this strategy, the seller's best response is to implement the buyer's first-best. In a bargaining framework, this equilibrium corresponds to a setting in which the buyer has full bargaining power. Analogously, there exist other equilibria in which the buyer's welfare $U(s_0)$ is in between the buyer's welfare in the equilibrium in this Proposition and the buyer's welfare under his first-best. A feature of these equilibria that might be considered less desirable is that the buyer's best response to a deviation by the seller depends on the buyer's belief whether the seller will return to his equilibrium strategy in the future. The equilibrium in this Proposition does not have this property, because the seller does not find it profitable to deviate by supplying less than $\breve{q}$ at any time before the limit-pricing phase, and because the buyer's indifference condition is binding in the limit-pricing phase. There also exist mixed-strategy variations of the equilibrium in Proposition 4.2 in which the buyer invests with a probability $\sigma \in (0,1)$ if $q < \breve{q}$: the seller's extra profits from supplying less than $\breve{q}$ in a near-zero interval if the buyer does not invest do not compensate for the long-run decrease in profits if the buyer does invest.

Thirdly, a larger share of the resource stock is sold at the buyer's reservation price, that is, $\breve{q}\left(\breve{T}-\breve{\Delta}\right) < q^*\left(T-\Delta\right)$.

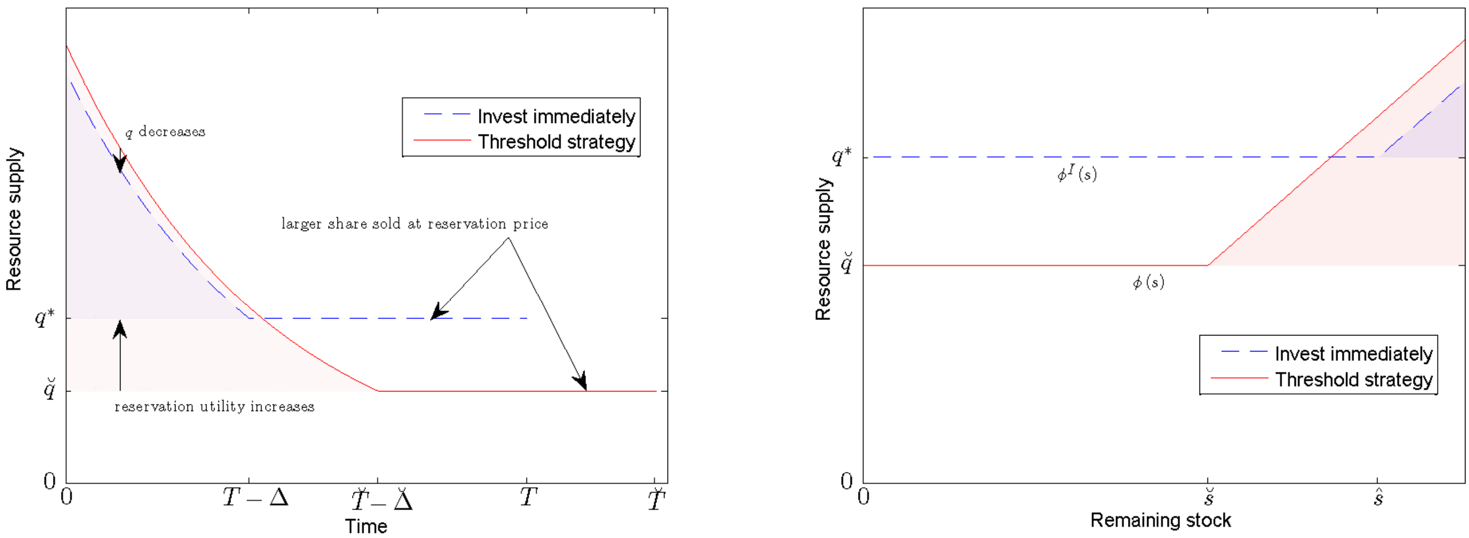

Figure 4.3: Resource supply before and after investment as a function of time (left) and the remaining stock (right)

Corollary 4.1. *$W(s) \downarrow 0$ when $r \downarrow 0$*

The buyer receives a positive surplus from the resource when the seller is too impatient to sell the entire stock at the buyer's reservation price. As r approaches zero, this consideration vanishes and so does $W(s)$.

In the Appendix, we derive the same results when the buyer can invest a continuous amount in each period and the marginal cost of the substitute depends on cumulative investment. The buyer invests the amount that maximizes consumption utility minus the interest on the investment costs in one go when the resource is exhausted. Similarly, the seller supplies at least the buyer's long-run utility until exhaustion.

4.6. Conclusion

Strategic exhaustible resource buyers have an incentive to delay the adoption of a substitute in order to depress current prices. When the buyer cannot commit to the arrival time of the substitute, a monopolistic seller can induce the buyer to delay adoption until exhaustion. In the absence of uncertainty or a fixed time-to-build delay, the threat of a substitute has a similar effect as an already available substitute in closed-loop equilibrium. The model is highly stylized - for example, OPEC and industrialized economies both face substantial coordination and free-riding problems that impede their ability to act as a monopolist or jointly develop a substitute.

Nonetheless, the results accord oil and gas exporters' efforts to stabilize prices, as unconventional hydrocarbons present an increasing threat to their revenues. In the European natural gas market, the emergence of shale gas has reduced Russia's monopoly power and its ability to use natural gas prices as a political weapon.[52] For the oil market, the Financial Times notes that "Saudi Arabia could respond [..] to the current US shale oil boom [by] boosting output and sending prices lower to stop the industry's development."[53]

4.A. Appendix

4.A.1. Proofs

Proof of Proposition 4.1

The seller supplies q^* at time t when the MR of supplying an additional unit is lower than the discounted MR at the time of exhaustion, i.e.

$$\pi'(q^*) \leq \lambda_t = \lambda_T e^{-r(T-t)} = ce^{-r(T-t)}$$

Solving for $T-t$ yields

$$T - t \leq -\frac{1}{r} \ln\left(\frac{\pi'(q^*)}{c}\right) \tag{4.9}$$

The seller supplies q^* in each period until exhaustion when this inequality holds for all $t \in \left(0, \frac{s}{q^*}\right)$. We can calculate the threshold stock s^* such that the seller practices limit

[52]"Nekrassov [a former Kremlin adviser] rejected suggestions that Russia might hit back by cutting off gas supplies, a tactic the country used in 2009 after the collapse of talks with Ukraine to end a row over unpaid bills and energy pricing. "Gas is no longer a weapon," Nekrassov said. "When Russia did that before, it realised that the foreign energy lobby reacted and efforts to find alternative sources were increased. If Russia kept threatening, it knows that nobody would be buying its gas in 20 years' time."" http://www.guardian.co.uk/world/2013/mar/23/cyprus-bailout-kremlin-reprisal-bank-levy, accessed April 2, 2013.

[53]http://www.ft.com/intl/cms/s/0/98d7f11c-43aa-11e2-844c-00144feabdc0.html#axzz2PJZKqbk8, accessed April 2, 2013.

pricing until exhaustion for all $s \leq s^*$ by substituting $T = \frac{s}{q^*}$ and solving for the s such that (4.9) holds with equality for $t = 0$

$$s^* = -\frac{q^*}{r} \ln \left(\frac{\pi' (q^*)}{c} \right)$$

When $s > s^*$, the seller extracts the last s^* units in a period of length $\Delta = \frac{s^*}{q^*}$. While extracting the first $s - s^*$ units, marginal revenue rises at the rate of interest and is equal the present value of marginal revenue at the time of exhaustion

$$\pi' (q_t) = ce^{-r(T-t)}$$

$$\Rightarrow q_t = \pi'^{-1} \left(ce^{-r(T-t)} \right)$$

Let $\eta(.) = (\pi')^{-1}$ be the seller's supply as a function of marginal revenue. The stock constraint determines T

$$\int_0^{T-\Delta} \eta \left(ce^{-r(T-t)} \right) dt + \Delta q^* = s \tag{4.10}$$

Proof of Proposition 4.2

We make use of a special case of Proposition 3 in (Hoel, 1983).

Lemma 4.1 (post-investment phase, (Hoel, 1983)). *When $s > s^*$, $\phi^I (s; c)$ is increasing in c.*

Because of the buyer's investment option, the resource is exhausted in finite time. This means there exists a subgame-perfect equilibrium in pure strategies. We assume that the number of periods N is sufficiently large, such that the optimal strategies are not affected by the length of the game. The buyer's strategy maximizes $U(s)$ with respect to k. His value from investing is

$$\begin{aligned}
U^I (s) &= \int_{T-\Delta}^{\infty} \bar{u} e^{-rt} dt - I + \int_0^{T-\Delta} u (q_t) e^{-rt} dt \\
&= \int_0^{\infty} (\bar{u} - rI) e^{-rt} dt + \int_0^T [u (q_t) - \bar{u}] e^{-rt} dt, \;\; q_t \text{ s.t. } u (q_t) \geq \bar{u} \; \forall t \in (0, T) \\
&= \int_0^{\infty} (\bar{u} - rI) e^{-rt} dt + W^I (s; c)
\end{aligned}$$

Let $\breve{U}(s)$ denote the payoff when the buyer adopts the equilibrium strategy, and denote the corresponding exhaustion time by $\breve{T}$. We have

$$\begin{aligned}
\breve{U}(s) &= \int_{\breve{T}}^{\infty} \bar{u}e^{-rt}dt - e^{-r\breve{T}}I + \int_0^{\breve{T}} u(q_t)e^{-rt}dt \\
&= \int_{\breve{T}}^{\infty} (\bar{u} - rI)e^{-rt}dt + \int_0^{\breve{T}} u(q_t)e^{-rt} \\
&= \int_0^{\infty} (\bar{u} - rI)e^{-rt}dt + \int_0^{\breve{T}} (u(q_t) - [\bar{u} - rI])e^{-rt}dt,\ q_t \text{ s.t. } u(q_t) \geq \bar{u} - rI\ \forall t \in \left(0, \breve{T}\right)
\end{aligned}$$

Then there exists a $\breve{c} \geq c$ such that

$$\breve{U}(s) = \int_0^{\infty} (\bar{u} - rI)e^{-rt}dt + W^I(s; \breve{c})$$

In order to prove that $\breve{U}(s) \geq U^I(s)$, it is sufficient to show that $W^I(s;c)$ is weakly increasing in c. By Lemma 4.1, $\phi^I(s;c)$ increases in c for $s > s^*$. Moreover, $\bar{u}$ decreases in c. Then $T - \Delta$ increases in c. Therefore, $W^I(s;c)$ is strictly increasing in c when $s > s^*$, that is, when the limit-pricing phase has not yet started. When $s \leq s^*$, observe that $\frac{\partial q^*}{\partial c} < 0$ and $\frac{\partial \Delta}{\partial c} \leq 0$, so $\frac{\partial s^*}{\partial c} \leq 0$. When the seller is limit pricing, he will keep limit-pricing after a reduction in c. Thus, $\frac{\partial W^I(s;c)}{\partial c} = 0$ for $s \leq s^*$.

The seller always supplies at least $\breve{q}$ given that the buyer only invests when $q < \breve{q}$. When the seller chooses a $q_t < \breve{q}$, he triggers immediate investment and is subject to the more stringent constraint $q_t \geq q^*$ for the remainder of the game. Choosing $q_t < \breve{q}$ therefore cannot be optimal, provided the period length ε is sufficiently small.

4.A.2. Continuous investment

In this section, we generalize the buyer's problem: rather than making a discrete investment decision, the buyer can invest a continuous amount in each period. The marginal cost of the substitute is a decreasing function of cumulative investment

$$c = c(I),\ c'(I) \leq 0,\ I = \int_0^{\infty} idt$$

We can write $\bar{u}(I) \equiv u(\psi^{-1}(c(I)))$. The timing in each stage is

(1) the seller supplies a quantity q_{t_i}

(2) the buyer chooses i

(3) the market clears at price $p_{t_i} = \min\left(\psi(q_{t_i}), c\left(I\right)\right)$

Profits and utility depend on the stance of technology I in addition to the supply q. Let $\phi\left(s, I\right)$ denote the seller's supply strategy and $\iota\left(s, I, q\right)$ the buyer's investment strategy. The value functions are given by

$$V\left(s, I\right) = \max_{\{q\}} \left\{\varepsilon\pi\left(I, q\right) + e^{-\epsilon r} V\left(s - \varepsilon q, I + \varepsilon\iota\left(s, I, q\right)\right)\right\} \tag{4.11}$$

$$U\left(s, I\right) = \max_{\{i\}} \left\{\varepsilon\left(u\left(I, \phi\left(s, I\right)\right) - i\right) + e^{-\varepsilon r} U\left(s - \varepsilon\phi\left(s, I\right), I + \varepsilon i\right)\right\} \tag{4.12}$$

Analogously to the main text, we can express the value of the resource to the buyer as

$$W\left(s, I\right) = U\left(s, I\right) - \left(\frac{\bar{u}\left(I\right)}{r} - I\right)$$

Utility from the substitute is a concave function of cumulative investment.

Assumption 4.1. $\dfrac{\partial \bar{u}}{\partial I} \geq 0$, $\dfrac{\partial^2 \bar{u}}{\partial I} < 0$

The equilibrium is similar to the one in Proposition 4.2. The buyer selects the cumulative investment level I^* that yields the highest long-run utility $\bar{u} - rI$, and invest this amount in one go when the resource is exhausted.

Proposition 4.3. *Let I^* be the solution to $\frac{\partial \bar{u}}{\partial I} = r$. There exists a subgame-perfect equilibrium in pure strategies. The buyer's equilibrium strategy is*

$$\iota\left(s, 0, q\right) = I^* \ \forall q : u\left(q\right) < \bar{u}\left(c\left(I^*\right)\right) - rI^* \tag{4.13a}$$

$$\iota\left(s, I, q\right) = 0 \ \ o.w. \tag{4.13b}$$

and the seller's equilibrium strategy is

$$\phi\left(s, I\right) = \phi^I\left(s; \min\left(c\left(I\right), u^{-1}\left(\bar{u}\left(c\left(I^*\right)\right) - rI^*\right)\right)\right) \tag{4.14}$$

with $\phi^I\left(s; c\right)$ as defined in Proposition 4.1.

Proof. First, we show that given the seller's strategy (4.14), the buyer's strategy must satisfy $\vec{I} \equiv \int_0^\infty i dt = I^*$. Suppose that for a certain buyer's strategy $\vec{I} < I^*$. Then by

Assumption 4.1, the buyer can marginally improve his welfare by choosing $\iota\left(0, \vec{I}, 0\right) > 0$. Suppose that $\vec{I} > I^*$. Then the buyer must either invest a positive amount when cumulative investment already exceeds I^*, or invest a large amount when $I < I^*$ such that cumulative investment overshoots I^*. We discuss both cases in turn. Firstly, suppose there exists an I such that $I^* \leq I$ and $\iota(s, I, \phi(s, I)) > 0$ for some s. Let I' denote the largest such I. Then the buyer can marginally improve his welfare by choosing $\iota(s, I', \phi(s, I')) = 0$. By Assumption 4.1, this increases the long-run utility $\bar{u} - rI$, and by Proposition 4.2, it increases the value of the resource $W(s, I)$. Secondly, suppose there exists an $I < I^*$ such that $\iota(s, I, \phi(s, I)) > I^* - I$ for some s. Let I'' be the largest such I. Then the buyer can marginally improve his welfare by choosing $\iota(s, I'', \phi(s, I'')) = I^* - I''$. Again, this improves long-run utility by Assumption 4.1 and increases the value of the resource by Proposition 4.2. By induction it follows that cumulative investment $\vec{I}$ must equal I^*.

Secondly, we show that given the seller's strategy (4.14), it is optimal to invest I^* in one go when the resource is exhausted. When the resource is exhausted, it is optimal for the buyer to reach the long-run optimal level I^* as quickly as possible. Suppose that the buyer's strategy entails $i_1 = \iota(0, I_1, 0) > 0$, $i_2 = \iota(0, I_2, 0) > 0$ for some $0 < I_1 < I_2 < I^*$. Then by Assumption 4.1, the buyer can marginally improve his welfare by choosing $\iota(0, I_1, 0) = I_2 - I_1 + i_2$. By induction it follows that the buyer invests $I^* - I$ when the resource is exhausted. When the resource is not yet exhausted, it is optimal to delay investments until exhaustion. Suppose there exists an $s > 0$ such that $\iota(s, I, \phi(s, I)) > 0$ for some I. Let s_3 be the lowest such s. Then the buyer can marginally improve his welfare by choosing $\iota(s_3, I, \phi(s_3, I)) = 0$ and increasing the investment at exhaustion by the same amount. This does not affect the long-run utility $\bar{u}(c(I^*))$, but increases $W(s, I)$ by Proposition 4.2.

Thirdly, the proof that the seller's strategy (4.14) is optimal given buyer's strategy (4.13) is analogous to the last part of the proof of Proposition 4.2. Q.E.D.

4.A.3. Commitment and closed-loop equilibria

In this section, we provide an example with an iso-elastic demand function in which a monopsonistic buyer that already has developed a substitute, as in Liski and Montero (2011), chooses a lower initial level of resource consumption under commitment than in closed-loop equilibrium. In the bilateral monopoly framework that we study in this paper, initial consumption is higher when the buyer can commit the invention time than in closed-loop equilibrium.

Let $u(q) = \sqrt{q}$, which corresponds to an iso-elastic demand with elasticity 2. We numerically compute the resource consumption paths in Liski and Montero's framework when the buyer can fully commit q_t for all t, in closed-loop equilibrium (where the buyer's strategy $q_t = C(S_t)$ is a function of the remaining stock) and in the social optimum. When the buyer has full commitment power, his problem is

$$\begin{aligned} V_{t=0} &= \max_{\{q_t,T\}} \int_0^T \{\sqrt{q_t} - p_t q_t\} e^{-rt} dt + \frac{1}{r}\frac{1}{4c} e^{-rT} \\ \text{s.t. } \dot{S}_t &= -q_t\, s_0 > 0,\ S_T = 0 \\ \dot{p}_t &= r p_t,\ p_T = c \end{aligned}$$

where T denotes the time at which the resource is exhausted and the buyer switches to the substitute. Liski and Montero, pp. 6 show that the buyer's commitment optimum is fully characterized by

$$\frac{1}{2\sqrt{q}} + \frac{1}{4r} q^{-1\frac{1}{2}} \frac{\partial q}{\partial t} = 0 \tag{4.15a}$$

$$\frac{1}{4c} - \frac{1}{2}\sqrt{q_T} = c r s_0 \tag{4.15b}$$

In closed-loop equilibrium, the buyer's value function is

$$V(S_t) = \int_t^\infty \left[\sqrt{C(S_t)} - P(S_t)\, C(S_t)\right] e^{-r(\tau - t)} d\tau$$

where $q_t = C(S_t)$ and $p_t = P(S_t)$ are the equilibrium consumption and pricing rule, respectively. The authors demonstrate (pp. 9) that the equilibrium is given by a first-order ODE in the pricing rule

$$\frac{1}{2}\sqrt{-\frac{P'(S)}{rP(S)}} - P(S) + P'(S)\, S = 0 \tag{4.16}$$

with boundary condition $P(0) = c$. Lastly, the social optimum is characterized by the Hotelling rule, $p_t = u'(q_t)$ and $p_T = c$ and has an analytical solution: $T =$

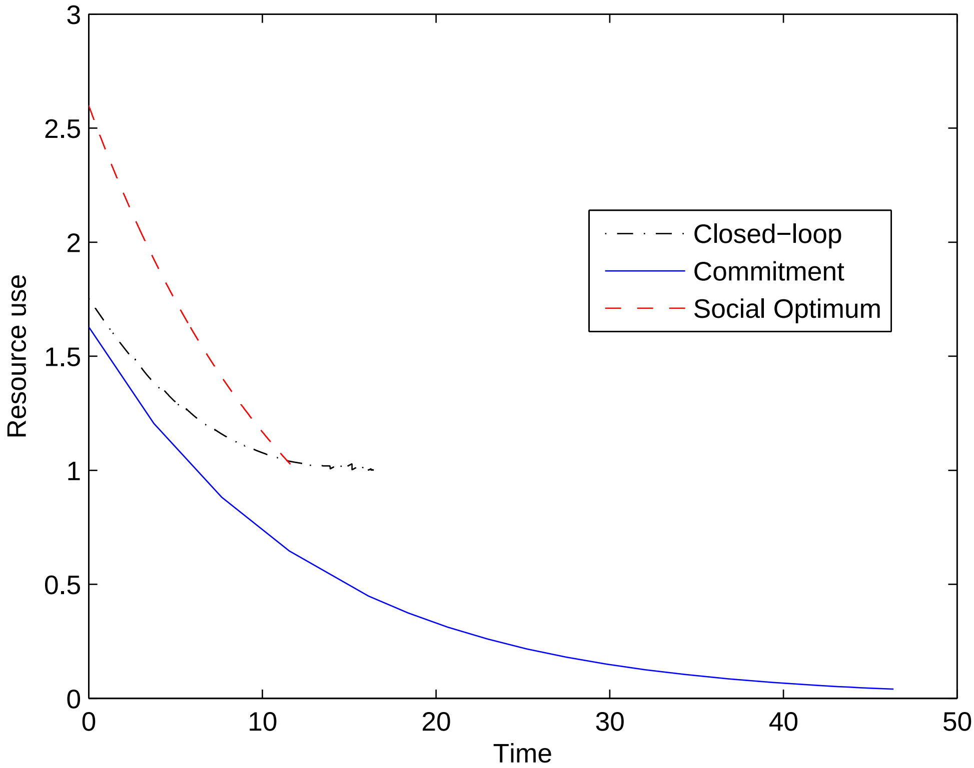

Figure 4.4: Resource consumption in Liski and Montero (2011)

$\frac{1}{2r}\ln(1+8rs_0c^2)$. Using MATLAB, we numerically solve (4.15) and (4.16) to determine the consumption paths under commitment and in closed-loop equilibrium for $s_0 = 20, c = 0.5, r = 0.04$.[54] Figure 4.4 depicts the results. In line with the intuition in the main text, resource use in Liski and Montero (2011) is lower under commitment than in closed-loop equilibrium. When the buyer has full commitment power, $p_t = p_T e^{-r(T-t)}$. By committing to a late switch to the substitute, the buyer reduces initial extraction but obtains a substantial price discount.

Now return to the bilateral monopoly model in this paper. Dasgupta et al. (1983) and Olsen (1986) discuss the buyer's optimal invention time under commitment at length. We demonstrate the following result through a series of Lemmas.

Proposition 4.4. *For iso-elastic demand with elasticity equal or larger than 2 and a sufficiently large initial stock, initial resource use is higher when the buyer commits the invention time at* $t = 0$ *than in closed-loop equilibrium.*

[54]The code is available upon request.

Lemma 4.2 (Proposition 1, Olsen (1986))**.** *For iso-elastic demand with elasticity equal or greater than 2 and* s_0 *sufficiently large, the seller has zero stock remaining at the buyer's optimal invention time.*

Lemma 4.3 (Dasgupta et al. (1983), pp. 1442)**.** *Resource supply at* $t = 0$ *increases in the invention time between 0 and the buyer's optimal invention time.*

Lemma 4.4 (Dasgupta et al. (1983), pp. 1442)**.** *For iso-elastic demand with elasticity equal or greater than 2 and* s_0 *sufficiently large, resource supply at* $t = 0$ *decreases in the buyer's invention time when the invention time is greater than the buyer's optimum.*

Lemma 4.1 implies that resource supply at $t = 0$ when the buyer commits to an infinitely large invention time (which is equivalent to $c \to \infty$) is larger than initial resource supply in closed-loop equilibrium. Then Lemmas 4.2, 4.3 and 4.4 establish the Proposition.

Chapter 5

Environmental catastrophes under time-inconsistent preferences[55]

5.1. Introduction

Many important ecosystems are subject to threshold dynamics: they can rapidly and irreversibly deteriorate when their vitality drops below a critical value. Lakes switch from a clear to a turbid state when the concentration of algae reaches a tipping point (Scheffer, 1997). Droughts, forest fires and logging may fuel a self-reinforcing replacement of tropical rainforest by grasslands in the Amazon (Nepstad et al., 2008). Ecologists hypothesize that species support ecosystem stability like rivets support a complex machine: initial component extractions do not affect the system's performance, but even a small number of further removals can trigger a sudden collapse (Ehrlich and Ehrlich, 1981). On a global scale, the climate system is subject to positive feedback mechanisms: the melting of polar ice caps will increase solar radiation absorption and permafrost melting in the Arctic could cause large methane releases (Lenton et al., 2008).

The threshold locations that govern these 'catastrophes' are highly uncertain, because of our limited knowledge of ecosystem behaviour and since current levels of environmental stress are without precedent (Muradian, 2001). This uncertainty poses an important economic tradeoff. Increasing our natural resource use yields temporary (using a piece of tropical wood in construction or burning a unit of fossil fuel) and/or permanent (bringing virgin land into production) economic benefits if we stay below the catastrophe thresholds, but incurs large and long-lasting damages if we do not. The consequences of temperature rises in the high single digits and upwards for example are likely to include large permanent loss of biodiversity, sea level rise and increased prevalence of

[55] This chapter also circulates as Michielsen (2013b).

extreme weather events. Because of their largely irreversible nature, the possibility of environmental catastrophes has important implications for intergenerational welfare analysis (for climate change, see e.g. Keller et al. (2004); Weitzman (2009, 2010)). This paper asks how concerns for catastrophe prevention affect the long-run concentration of pollutants and the allocation of natural resource use across generations.

To answer this question, I use a welfare criterion that balances both present and far-distant future outcomes.[56] The welfare of generation t is a weighted sum of expected discounted utility and the probability that an irreversible catastrophe will occur at any point in the future

$$W(t) = \int_t^\infty \mathbb{E}\, u(s)\, e^{-\delta(s-t)} ds - \xi P[\tau < \infty]$$

where τ is the occurrence time of the catastrophe. The welfare function generates a time-inconsistency: the current generation would like to sacrifice their descendants' consumption for the long-run objective, but the descendants themselves are not as willing to make these sacrifices once they inherit the economy. When the current generation recognizes that future generations have different preferences, its response depends on the nature of the environmental problem. If the pollutant that causes the catastrophe risk is expected to become obsolete in the near future or if the risk is related to emissions from a scarce exhaustible resource, the current generation may reduce its consumption in an attempt reduce the maximum pressure on the ecosystem and hence avert a catastrophe. If the pollutant is abundant and expected to remain essential, the catastrophe becomes a self-fulfilling belief.

The literature on optimal control in environmental problems under time-inconsistent preferences is scarce. Li and Löfgren (2000) look at renewable resource management with similar preferences as in the present paper, but restrict themselves to full commitment and thus assume away the time-inconsistency problem. Karp (2005) and Gerlagh and Liski (2012) study Markov-perfect climate mitigation strategies when regulators use hyperbolic discounting. An important difference with the present paper is that generations

[56] Chichilnisky (1996), Alvarez-Cuadrado and Long (2009) and Long and Martinet (2012) propose related welfare functions. Chichilnisky (1996) discusses a criterion that consists of a weighted sum of discounted utility and lim-inf utility. Alvarez-Cuadrado and Long (2009) advocate a weighted sum of discounted utility and a Rawlsian maxi-min criterion; Long and Martinet (2012) propose a weighted sum of discounted utility and an endogenous set of minimum rights to be guaranteed to all generations.

with hyperbolic preferences do not explicitly care about the distant future; they merely place a higher weight on their own felicity. This feature causes hyperbolic regulators with full commitment power to stabilize emission stocks at a lower level, but start off with higher emission flows than in Markov equilibrium (Karp, 2005). This ranking between the commitment and Markov solutions does not always hold with the preferences in the present paper - specifically, it breaks down in a model that is close to Karp (2005).

Karp and Tsur (2011) consider catastrophic climate change under hyperbolic preferences in a discrete-choice setting. Mitigation decisions are strategic complements across generations, and perpetual stabilization and perpetual business-as-usual can both be Markov equilibria. Different from the present paper, the catastrophe hazard in Karp and Tsur (2011) persists even when emissions cease perpetually: emissions irreversibly increase the hazard in all future periods, but do not affect not the catastrophe hazard in the current period. The range of equilibria in Karp and Tsur (2011) is sensitive to the functional form of the hazard rate. In equilibrium, generations can only cease emissions at concentration levels at which additional emissions increase the hazard sufficiently strongly, because large increases in the hazard deter future generations from reneging on the current generation's plan to stabilize the carbon concentration.[57]

It is difficult to infer long-term preferences for environmental goods from market data. There is a dearth of investment assets with very long horizons, and extrapolating preferences from shorter-term decisions requires contentious assumptions. Nordhaus (1994) argues that revealed preferences in the capital market indicate a high degree of impatience. He calibrates a Ramsey discount rate of an infinitely-lived agent that uses exponential discounting, and finds a pure rate of time preference of 3% - implying negligible welfare weights beyond a 50-year horizon. His result is sensitive to both the infinitely-lived agent and the exponential discounting assumptions. Observed saving decisions are consistent with concerns for the medium or distant future if we consider a different preference structure, for example that individuals discount consumption within their own lifetime but not across generations (Dasgupta, 2012) or hyperbolic discounting (Gerlagh and Liski, 2012).

Stated-preference studies circumvent this problem and find that people care about

[57]For stationary optimal control under catastrophic risk, see e.g. Cropper (1979); Reed (1984); Tsur and Zemel (1996, 2008); Polasky et al. (2011). Contrary to some of these papers, there is no endogenous (effective) discounting in the present work.

long-term environmental outcomes, consonant with my welfare criterion. Layton and Levine (2003) calibrate an exponential discounting model and estimate a 0.7% median discount rate for climate mitigation measures, whereas Layton and Brown (2000) find no appreciable difference in willingness to pay for environmental damages that occur in 60 or 150 years. Gattig and Hendrickx (2007) survey evidence that non-monetary indicators of the perceived severity of environmental risks, such as the willingness to engage in pro-environmental behaviours, are unresponsive to the temporal delay of environmental impacts. The catastrophe term in my welfare function also captures the nonuse value of natural assets, which may constitute more than half of their total economic value (Greenley et al., 1981; Kaoru, 1993; Langford et al., 1998; Wattage and Mardle, 2008). A large part of the value people attach to preserving the environment is not related to current or future use, but to simply knowing that a species or pristine area exists. When the value of e.g. species protection does not depend on current and future use, the welfare loss from future extinctions is likely independent of the time of occurrence.

My welfare criterion also addresses deontological motives. The Lockean proviso states that appropriating natural resources for current use is justified only if 'enough and as good' is left for the future. Within a purely consequentialist framework, the risk of a future catastrophe can be offset by an increase in current consumption. Even if future generations are compensated, they cannot consent to any compensation. The second term in the welfare function reflects the difficulty of compensating future generations for the loss of vital ecosystems, and the uncertainty whether they would be willing to accept an increase in man-made goods in return. Lastly, the catastrophe term captures an intrinsic aversion to the idea that the human community, encompassing both current and future generations, will at some point cause an environmental catastrophe.

The widely-used utilitarian criterion (Nordhaus, 1994; Stern, 2007) results in either a 'dictatorship of the present' or a 'dictatorship of the future' (Chichilnisky, 1996). With a zero discount rate, the utilitarian approach is insensitive to near-term outcomes, because the generations that are alive today are vastly outnumbered by their far-future counterparts. With a positive discount rate, the utilitarianist attaches near-zero weight to the distant future, as its importance is diminished by compounded discounting.[58]

[58] Weitzman (2009) shows that the present value of expected losses from future catastrophes may be infinitely large even with a positive discount rate, but his assumptions have been subject to much critique (Millner, 2013). Most prominently, his result requires the utility function to be unbounded from

Table 5.1: Time inconsistent welfare weights

	current utility	future utility	catastrophe
current generation	1	$\rho < 1$	ξ
future generation		1	ξ

These properties also apply to hyperbolic preferences, depending on whether the long-term discount rate is zero or not - but not to the preferences in this paper.

Under the proposed criterion with an explicit concern for catastrophe prevention, I demonstrate how optimal resource use depends on the nature of the environmental problem. I consider a sequence of models with a common framework. A series of non-overlapping generations derive utility from an emission-intensive consumption good. Emissions from production add to a pollution stock. In each period, a constant fraction of the stock decays naturally.[59] A catastrophe occurs when the pollution stock exceeds an unknown threshold. The risk can be eliminated by keeping the stock at or below its current level, which is known to be 'safe'.[60] Importantly, each generation's intrinsic welfare loss from a catastrophe does not depend on the time of occurrence.[61] Table 5.1 illustrates the inconsistency: the current generation discounts future utility relative to its intrinsic welfare loss from a catastrophe, but future generations do not discount their own utility relative to the catastrophe loss. As a consequence, future generations emit too much from the current generation's perspective and a dynamic game ensues. Generations have a strategic motive to distort their emissions in order to influence future emissions. I compare emissions and the probability of a catastrophe in three cases: (a) when the first generation can commit all current and future emissions (the commitment solution), (b) when current generations do not anticipate that future generations have different preferences (the naive solution) and (c) when current generations take into account the reaction of future generations (the Markov equilibrium).

I firstly introduce a two-period model. This model represents a setting in which the

below. See Buchholz and Schymura (2012) for an elaborate discussion.

[59]In a broader interpretation, we may think of the emission flows as the one-off benefits of bringing additional natural resources under cultivation (e.g. cutting down a forest), and the natural decay as the flow of benefits that cultivated resources can sustainably provide (such as agricultural products).

[60]This type of catastrophe risk is also studied in Tsur and Zemel (1994, 1996); Nævdal (2006).

[61]I explicitly allow for the possibility that the catastrophe also has a direct effect on utility.

catastrophe risk is expected to recede in the near future, for example because technological change will make the polluting resource obsolete. The first generation may be more or less cautious under commitment than in Markov equilibrium, depending on the utility and threshold distribution functions. Because the number of future generations that can affect the catastrophe risk is small, the current generation has a direct influence on future decisions. When current and future emissions are strategic substitutes, today's generation can pass on the costs of catastrophe prevention to the future by increasing its emissions. I derive unambiguous results for two functional forms.

Secondly, I consider an infinite-horizon model with an abundant pollutant. This model is informative when the pollutant is plentifully available and will remain essential for a long period. Reserves of coal are sufficient to last another 200 years and pose a serious threat to the global climate unless we develop a substitute. We may also interpret the pollution stock as the total amount of deforested land: the pressure to convert rainforests for agricultural use is unlikely to let up any time soon. In Markov equilibrium, the catastrophe becomes a self-fulfilling prophecy. The steady-state pollution stock depends on beliefs. Given consistent beliefs, individual generations cannot influence the steady state, and will conclude that mitigation efforts are futile. There even exists an equilibrium in which the degree of catastrophe aversion has no effect on equilibrium behaviour, that is, generations act as if they do not care about the long-run future. As opposed to under hyperbolic preferences as in Karp (2005) and Gerlagh and Liski (2012), who also employ infinite-horizon models with abundant pollutants, not only steady-state emission stocks but also emission flows are higher in Markov equilibrium than under commitment. Naive policies also lead to high pollution stocks eventually, but degrade the environment less rapaciously. Because naive generations mistakenly believe that pollution concentrations can be stabilized at a low level, they choose lower emissions than under full rationality.

Lastly, I propose an infinite-horizon model with a scarce pollutant, which is relevant for local pollution problems related to exhaustible resource extraction. The pollution stock first increases, but later declines when reserves of the resource become depleted. When the initial resource reserve is sufficiently small, future generations have limited ability to increase the pollution stock. Early generations then have an incentive to reduce emissions in Markov equilibrium that is not present under commitment. By reducing their own resource use, early generations smooth the time path of emissions, allowing natural decay to reduce the maximum pollution stock and hence the probability of a catastrophe. I provide a numerical example in which initial emissions in Markov equilibrium are lower than under commitment.

The results from the infinite-horizon model with an abundant pollutant offer an explanation why climate change mitigation efforts are far below the level necessary to limit temperature increases to two degrees. The embodied carbon in global reserves of coal and unconventional oil and gas exceeds cumulative historical emissions by a multiple (Kharecha and Hansen, 2008), and natural carbon sinks are insufficient to stabilize the concentration in the atmosphere unless emissions decrease significantly. Dangerous climate change will not be averted because of fossil fuel scarcity or carbon dissipation; only by deliberate and costly reductions in fossil fuel consumption. Rational policymakers who are not willing to foot the bill for the long-term objective of limiting climate change recognize that their successors are also unwilling to pay. Because the objective of stabilization at relatively low concentration levels is out of reach, inaction becomes a self-fulfilling equilibrium.

5.2. Two-period model

Consider a model with two generations, living in periods $t = 1, 2$. A representative agent in each generation derives utility $u_t(z_t)$ from an emission-intensive consumption good z, the economy's single commodity (hereafter: emissions). The utility functions satisfy $u_t' \geq 0$, $u_t'' \leq 0$, $\exists\ \bar{u}_t : u_t(z) < \bar{u}_t\ \forall\ z$. Emissions z_t contribute to a pollution stock D_t. Natural decay is relatively unimportant when the number of time periods is small, so I abstract from it in this model. I normalize $D_0 = 0$.

$$D_t = D_{t-1} + z_t,\ D_0 = 0$$

A catastrophe occurs when the stock reaches an unknown threshold $\hat{D}$. The threshold is randomly distributed on the interval $\left[0, \bar{D}\right]$. I express the probability of a catastrophe as a function of cumulative emissions through pdf $f(D)$ and cdf $F(D)$.

Each generation's (ex ante) welfare w_t is given by a weighted sum of discounted utility (positive) and the probability that a catastrophe will occur in either period (negative).

The first generation discounts utility of the second generation by a factor $\rho < 1$, but its welfare loss from a catastrophe does not depend on the time of occurrence. I distinguish between three cases. If the threshold is never breached ($D_2 < \hat{D}$), we disregard the catastrophe term in the welfare functions and ex post welfare W_t is

$$\begin{aligned} W_1 &= u_1(z_1) + \rho u_2(z_2) \\ W_2 &= u_2(z_2) \end{aligned}$$

If the threshold is breached in the second period ($D_1 < \hat{D} < D_2$), both generations suffer an intrinsic catastrophe welfare loss ξ:

$$\begin{aligned} W_1 &= u_1(z_1) + \rho u_2(z_2) - \xi \\ W_2 &= u_2(z_2) - \xi \end{aligned}$$

When the threshold is breached in the first period ($D_1 > \hat{D}$), the second generation receives utility $\underline{u}$, to capture the impacts of a catastrophe on material well-being.[62]

$$\begin{aligned} W_1 &= u_1(z_1) + \rho \underline{u} - \xi \\ W_2 &= \underline{u} - \xi \end{aligned}$$

The welfare functions for the two generations read

$$w_1 = u_1(z_1) + (1 - F(z_1))\,\rho u_2(z_2) + F(z_1)\,\rho \underline{u} - \xi F(z_1 + z_2) \tag{5.1a}$$

$$w_2 = \begin{cases} u_2(z_2) - \xi \frac{F(z_1+z_2) - F(z_1)}{1-F(z_1)} & \text{if } z_1 < \hat{D} \\ \underline{u} - \xi & \text{if } z_1 \geq \hat{D} \end{cases} \tag{5.1b}$$

The second generation observes whether the first generation's emissions have triggered the catastrophe or not,[63] so it evaluates catastrophe risk using the conditional cdf

[62] In the remainder of this paper, I assume $\underline{u} > -\infty$ to be sufficiently small such that the catastrophe is also undesirable from a point of view of utility maximization. This is not necessary for the formal analysis however. If the catastrophe does not affect utility, all post-catastrophe generations choose z_t arbitrarily large and $\underline{u} = \bar{u}_t$.

[63] When the catastrophe is only observed at the end of the second period, the second generation chooses a higher z_2 because there is a probability that the first generation has already triggered the catastrophe, in which case second-period mitigation is fruitless.

$\frac{F(z_1+z_2)-F(z_1)}{1-F(z_1)}$.[64] The discount factor generates time-inconsistency in the preference structure: the second generation places a higher weight on second-period utility $u_2(z_2)$ relative to the probability of a catastrophe $F(D_2)$ than the first generation does.

I distinguish between three solutions. Firstly, the commitment solution (superscript C), in which the first generation commits all current and future emissions. Secondly, the 'naive' solution (superscript N), in which the first generation does not anticipate that future generations will make a different trade-off between $u_2(z_2)$ and $F(D_2)$. Lastly, I consider the Markov solution (superscript M), in which the first generation foresees the preference reversal of the second generation and selects z_1 by backward induction, maximizing its welfare given the optimal response of the second generation.

5.2.1. Commitment solution

When the first generation can commit second-period emissions conditional on whether the threshold is breached in the first period, z_1^C and z_2^C immediately follow from (5.1a) in case of an interior solution

$$\underbrace{u_1'\left(z_1^C\right)}_{I}-\underbrace{\rho f\left(z_1^C\right)\left[u_2\left(z_2^C\right)-\underline{u}\right]}_{II}-\underbrace{\xi f\left(z_1^C+z_2^C\right)}_{III}=0 \tag{5.2a}$$

$$\rho u_2'\left(z_2^C\right)-\xi\frac{f\left(z_1^C+z_2^C\right)}{1-F\left(z_1^C\right)}=0 \quad \text{if } z_1^C<\hat{D} \tag{5.2b}$$

The first generation equates discounted marginal utility in both periods with the marginal welfare loss from catastrophe risk. The three components of (5.2a) represent the first generation's considerations. The first term is the first generation's marginal utility. The second term indicates that higher first-period emissions increase the probability of reducing second-period utility to $\underline{u}$. The third term reflects the first generation's intrinsic desire to prevent a catastrophe. When the welfare weight on catastrophe prevention is sufficiently low, we may have a corner solution and $\left(z_1^C, z_2^C\right) \to (\infty,\infty)$.

[64] When the first generation is ambiguity-averse, this Bayesian updating would also be a source of time inconsistency.

5.2.2. Naive solution

In the naive solution, the first generation behaves as if it could commit both z_1 and z_2. The second generation however selects z_2^N to maximize (5.1b) rather than (5.1a), yielding

$$u_1'\left(z_1^N\right) - \rho f\left(z_1^N\right)\left[u_2\left(z_2^C\right) - \underline{u}\right] - \xi f\left(z_1^N + z_2^C\right) = 0 \tag{5.3a}$$

$$u_2'\left(z_2^N\right) - \xi \frac{f\left(z_1^N + z_2^N\right)}{1 - F\left(z_1^N\right)} = 0 \quad \text{if } z_1^N < \hat{D} \tag{5.3b}$$

By definition, z_1 is the same in the naive solution as in the commitment solution. Substituting $z_1^N = z_1^C$ in (5.3b) and comparing with (5.2a), $z_2^N > z_2^C$: the second generation chooses higher second-period emissions than the first generation would have under commitment.[65]

5.2.3. Markov solution

In the Markov solution, the first generation correctly anticipates the second generation's reaction. Condition (5.3b) implicitly defines the second generation's reaction function $r\left(z_1\right)$

$$u_2'\left(r\left(z_1^M\right)\right) = \xi\, \frac{f\left(z_1^M + r\left(z_1^M\right)\right)}{1 - F\left(z_1^M\right)} \tag{5.4}$$

To avoid clutter, I omit the superscript M in the derivation of the reaction function. Differentiating with respect to z_1, I obtain

$$\begin{aligned} u_2''\left(r\left(z_1\right)\right) r'\left(z_1\right) &= \xi \left(\frac{f'\left(z_1 + r\left(z_1\right)\right)\left[1 - F\left(z_1\right)\right] + f\left(z_1\right) f\left(z_1 + r\left(z_1\right)\right)}{\left[1 - F\left(z_1\right)\right]^2} + \right. \\ &\qquad \left. r'\left(z_1\right) \frac{f'\left(z_1 + r\left(z_1\right)\right)}{1 - F\left(z_1\right)} \right) \\ \Leftrightarrow r'\left(z_1\right) &= \xi\, \frac{f'\left(z_1 + r\left(z_1\right)\right)\left[1 - F\left(z_1\right)\right] + f\left(z_1\right) f\left(z_1 + r\left(z_1\right)\right)}{\left[1 - F\left(z_1\right)\right]\left[u_2''\left(r\left(z_1\right)\right)\left[1 - F\left(z_1\right)\right] - \xi f'\left(z_1 + r\left(z_1\right)\right)\right]} \end{aligned} \tag{5.5}$$

The condition for the numerator in (5.5) to be positive is similar to f having an increasing hazard function. The sign of the denominator depends on the curvature of f. A sufficient condition for the second generation's reaction function to be downward-sloping is $f'\left(z_1 + r\left(z_1\right)\right) \geq 0$. When $f'\left(z_1 + r\left(z_1\right)\right)$ is sufficiently negative, an increase in z_1 low-

[65] In addition to a corner solution in both periods, we may now also have a corner solution in the second period only.

ers the marginal probability of a catastrophe to such an extent that it becomes attractive for the second generation to choose a higher emission level.

The first-order condition for the first generation is

$$u_1'(z_1) - \rho f(z_1) u_2(r(z_1)) + \rho [1 - F(z_1)] u_2'(r(z_1)) r'(z_1) + \rho f(z_1) \underline{u} -$$
$$\xi f(z_1 + r(z_1))(1 + r'(z_1)) = 0$$
$$\Leftrightarrow \underbrace{u_1'(z_1)}_{I} - \underbrace{\rho f(z_1)[u_2(r(z_1)) - \underline{u}]}_{II} - \xi \underbrace{(1-\rho) f(z_1 + r(z_1)) r'(z_1)}_{IV} - \xi \underbrace{f(z_1 + r(z_1))}_{III} = 0 \tag{5.6}$$

Terms I, II and III are also present in the commitment FOC and have the same interpretation. However, as I discussed in section 5.2.2, $r\left(z_1^C\right) > z_2^C$. The points at which terms II and III are evaluated are different than in the commitment solution. Holding z_1 constant, term II is unambiguously larger in the Markov solution: because the second generation chooses higher emissions, the utility loss to the second generation $u_2(z_2) - \underline{u}$ in case of a first-period catastrophe is larger than under commitment. This effect makes the first generation more cautious. Whether term III makes the first generation more conservationist in Markov equilibrium depends on the local curvature of the threshold pdf. The Markov FOC also contains an additional term IV that is not present in the commitment FOC. This is the strategic motive to influence the second generation's emissions through the second-period catastrophe hazard. When $r'(z_1)$ is negative (positive), the first generation can reduce z_2 by increasing (decreasing) its own emissions.

Comparing (5.6) and (5.2a), it is not possible to say whether first-period emissions are higher in the Markov or in the commitment solution without assuming functional forms for u_t and F. The Appendix contains two examples with different rankings of z_1^C and z_1^M.

When catastrophe risk is expected to recede in the medium term, current decision makers can directly influence their successors' actions and the probability of a catastrophe. Interestingly, the desire to reduce perceived 'overconsumption' by future generations can lead current decision makers to increase their own emissions, even if they so increase the probability of a catastrophe. The results from this section are less relevant when catastrophe risk persists over long horizons, for example because the pollutant remains essential into the far future: the current generation has limited ability to affect the policies of distant generations. The infinite-horizon models in the next two sections deal with more persistent risks.

5.3. Infinite horizon, abundant pollutant

Consider an infinite-horizon model with a continuum of non-overlapping generations and an abundant pollutant. As in the previous section, each generation derives utility from its own emissions $u\left(z\left(t\right)\right)$ and cares about future utility (discounted at rate δ) as well as the possibility of a catastrophe occurring at some point the future. A constant fraction α of the pollution stock decays in each period, so that

$$\dot{D} = z - \alpha D \tag{5.7}$$

Utility is concave and bounded. The pollution stock only has a direct effect on utility when a catastrophe occurs. The hazard rate $\psi\left(D\right) \equiv \frac{f(D)}{1-F(D)}$ is increasing.

Assumption 5.1. *$u\left(D, z\right) = u\left(z\right),\ u'\left(z\right) > 0,\ u''\left(z\right) < 0\ \forall z$ and $\lim_{z\to\infty} u\left(z\right) = \bar{u}$.*

Assumption 5.2. *$\psi'\left(D\right) \geq 0$.*

When the catastrophe occurs, all subsequent generations receive utility $\underline{u} < \bar{u}$. As in section 5.2, a catastrophe is immediately observable, and generations condition their strategy on whether the catastrophe has occurred already. Because the post-catastrophe game is trivial, I focus on pre-catastrophe strategies. Throughout, I assume existence of optimal solutions and that $D\left(t\right)$ is non-decreasing along the optimal path.[66] The intuition for this assumption is as follows. Keeping the stock constant already eliminates the catastrophe hazard. A trajectory in which the stock is V-shaped or declining during an interval of time results in lower discounted utility than an alternative path that keeps the stock constant over the same interval, without reducing the probability of a catastrophe.

Define $\eta\left(t\right) \equiv \psi\left(D\left(t\right)\right)\left(z\left(t\right) - \alpha D\left(t\right)\right)$ as the instantaneous catastrophe hazard at time t and $H\left(t\right) \equiv \int_0^t \eta\left(s\right) ds$ as the cumulative hazard, and let τ denote the occurrence

[66] Tsur and Zemel (1996) prove these properties for $\xi = 0$.

time of the catastrophe. For the remainder of this paper, W denotes a generation's welfare given a future emissions path, and V denote welfares at this generation's optimal decision. For a given admissible trajectory $z(s)$, the welfare of generation t is[67]

$$\begin{aligned} W^t(D(t)) = & \mathbb{E}\left(\int_t^\infty \left(u(z(s))\,\mathbf{1}_{\tau>s} + \underline{u}\mathbf{1}_{\tau\le s}\right)e^{-\delta(s-t)}ds\right) - \xi\,\frac{P[\tau\in[t,\infty)]}{1-P[\tau\in[0,t)]} \\ & s.t.\ \dot{D} = z-\alpha D,\ D(t)=D_t \\ = & \int_t^\infty \left(u(z(s))\left[1-(H(s)-H(t))\right] + \underline{u}\left[H(s)-H(t)\right]\right)e^{-\delta(s-t)}ds \\ & -\xi\,\frac{P[\tau\in[t,\infty)]}{1-P[\tau\in[0,t)]} \\ & s.t.\ \dot{D} = z-\alpha D,\ D(t)=D_t,\ \dot{H}=\psi(D)(z-\alpha D) \end{aligned} \tag{5.8}$$

In Appendix 5.A.3, I outline the necessary conditions for stationary dynamic optimization problems with uncertain thresholds, as derived in Nævdal (2006).

5.3.1. Commitment solution

If the first generation can commit all current and future emissions, it maximizes (5.8) for $t=0$. Its problem is

$$\begin{aligned} \max_z \Big\{ & W^C(D(0)) = \int_0^\infty \left(u(z(s))\left[1-H(s)\right]+\underline{u}H(s)\right)e^{-\delta s}ds - \xi\ P[\tau\in[0,\infty)] \\ & s.t.\ \dot{D}=z-\alpha D,\ D(0)=D_0,\ \dot{H}=\psi(D)(z-\alpha D)\Big\} \end{aligned} \tag{5.9}$$

I may rewrite the problem by including the intrinsic welfare loss from a catastrophe in the integral of utility.

$$\begin{aligned} \max_z \Big\{ & W^C(D(0)) = \int_0^\infty \left(u(z(s))\left[1-H(s)\right]+\underline{u}H(s)-\eta(s)\,\xi e^{\delta s}\right)e^{-\delta s}ds \\ & s.t.\ \dot{D}=z-\alpha D,\ D(0)=D_0,\ \dot{H}=\psi(D)(z-\alpha D)\Big\} \end{aligned} \tag{5.10}$$

The $e^{\delta s}$ term in the round brackets ensures that the intrinsic welfare loss from a catastrophe is constant in present terms, regardless of the time of occurrence. As time passes, it becomes prohibitively costly from the first generation's point of view to risk a catastro-

[67] τ is distributed as a Poisson process, as described in the Appendix. For brevity, I omit the distribution of τ in the main text.

phe, because the utility discount rate diminishes the benefits of future emissions relative to the intrinsic catastrophe loss. The first generation therefore stabilizes the emissions stock at some finite date t' such that the marginal benefit of increasing the pollution stock (higher current utility and higher steady-state utility if the threshold is not breached) equals the expected marginal cost (a permanent decrease in utility and the intrinsic welfare loss evaluated at $\tau = t'$ if the catastrophe does occur).

Proposition 5.1. *The commitment solution is characterized by a steady-state pollution stock D^C. There exists a $t' < \infty$ such that $D^C(t') = D^C$ and $z^C(t) = \alpha D^C \; \forall \; t \geq t'$. D^C and t' satisfy*

$$u'\left(\alpha D^C\right) = \frac{\psi\left(D^C\right)}{\delta + \alpha}\left[u\left(\alpha D^C\right) - \underline{u} + \delta \xi e^{\delta t'}\right] \tag{5.11}$$

A formal analysis of the comparative statics of the steady state is complicated by the presence of two endogenous variables in (5.11), D^C and t'. In the next subsections, I derive comparative statics for the naive and Markov steady states and discuss the intuition behind them.

5.3.2. Naive solution

In the naive solution, each generation t solves a problem that is similar to (5.10), with the initial pollution stock determined by previous generations.

$$\begin{aligned} \max_{z} \Bigg\{ & W^{t,N}\left(D\left(t\right)\right) = \int_t^\infty \left(u\left(z\left(s\right)\right)\left[1 - \left(H\left(s\right) - H\left(t\right)\right)\right] + \underline{u}\left[H\left(s\right) - H\left(t\right)\right] \right. \\ & \left. -\eta\left(s\right)\xi e^{\delta(s-t)}\right) e^{-\delta(s-t)} ds \\ & s.t.\; \dot{D} = z - \alpha D,\; D\left(t\right) = D_t,\; \dot{H} = \psi\left(D\right)\left(z - \alpha D\right) \Bigg\} \end{aligned} \tag{5.12}$$

In (5.12), the intrinsic welfare loss from triggering a catastrophe at time s evaluated by generation t decreases in t (ξ is multiplied with $e^{\delta(s-t)}$, compared to $e^{\delta s}$ in the commitment solution). Each generation t envisions a preferred steady-state stock $D^{t,N}$, but as every subsequent generation places a higher weight on its own utility, and thus a lower relative weight on catastrophe prevention, the stock targets $D^{t,N}$ increase over time. The targets converge to a unique level D^N that even the most distant generations do not want to exceed, as the marginal welfare gain of higher steady-state utility falls short of the permanent utility reduction and the welfare loss associated with a catastrophe.

Proposition 5.2. *The solution to generation t's problem is characterized by a steady-state stock $D^{t,N}$. Let D^N be given by*

$$u'\left(\alpha D^N\right) = \frac{\psi\left(D^N\right)}{\delta+\alpha}\left[u\left(\alpha D^N\right) - \underline{u} + \delta\xi\right] \tag{5.13}$$

Then

(i) $D^{t,N} < D^N \ \forall\ t$ *and* $\lim_{t\to\infty} D^{t,N} = D^N$

(ii) $\frac{\partial D^N}{\partial \alpha} \gtreqless 0$ *iff*

$$(\alpha+\delta)^2 D^N u''\left(\alpha D^N\right) - (\alpha+\delta) D^N \psi\left(D^N\right) u'\left(\alpha D^N\right) + \psi\left(D^N\right)\left(u\left(\alpha D^N\right) - \underline{u} + \delta\xi\right) \gtreqless 0$$

(iii) $\frac{\partial D^N}{\partial \delta} \gtreqless 0$ *iff* $u\left(\alpha D^N\right) - \underline{u} - \alpha\xi \gtreqless 0$

The left and right hand side of (5.13) represent the marginal benefit and cost of increasing the steady-state stock, respectively. Because of Assumptions 5.1 and 5.2, the left hand side is decreasing and the right hand side is increasing in D. Therefore, it cannot be optimal for any generation t that inherits stock D^N to choose $z^{t,N}(t) > \alpha D^N$. As a consequence, the pollution stock never exceeds D^N.

The net effect of the pollution decay rate α on the steady-state stock D^N is ambiguous. When α increases, a given stock level allows for higher emissions without risking a catastrophe. This effect increases the steady state stock. However, holding D^N constant, marginal utility $u'(\alpha D)$ decreases and the utility loss from a catastrophe $u\left(\alpha D^N\right) - \underline{u}$ increases, which depresses D^N. A higher discount rate δ also has two opposing effects. On one hand, it increases the relative weight of the current gain of increasing the stock $u'(\alpha D)$ compared to the stream of possible future utility reductions $(u(\alpha D) - \underline{u})/\delta$. This effect encourages higher steady-state stocks. On the other hand, it also increases the relative importance of the intrinsic catastrophe loss ξ compared to the stream of future utility gains if no catastrophe occurs. This consideration decreases D^N. The net effect depends on the relative magnitudes. If the utility loss from a catastrophe is small, the latter effect is more important. If the weight of the intrinsic catastrophe loss is small, the former effect dominates.

Corollary 5.1. $D^C \leq D^N$

The steady-state stock is higher in the naive solution than in the commitment solution. Future generations have higher relative welfare weights on their own utility, and thus reoptimize towards higher steady-state pollution stocks.

5.3.3. Markov solution

A Markov equilibrium is defined by a policy function $\zeta^M(D)$ such that $z^M(t) = \zeta^M(D(t))\ \forall\, t$.[68] In Proposition 5.3, I show that there exists a continuum of Markov equilibria which can be ranked by their steady-state pollution stocks. Early generations' emissions depend on their beliefs about future emissions. When generation t believes that future generations will increase the stock up to a certain level D^M, its choice of $z^M(t)$ has no effect on the maximum pollution stock. Each generation thus maximizes expected discounted utility subject to the stock not exceeding the perceived maximum. The range of equilibria is bounded by two considerations. The equilibrium steady-state stock cannot exceed the level that maximizes expected discounted utility (the first component of (5.8)) disregarding the intrinsic loss. When the perceived steady-state stock is below the naive steady-state D^N, far-future generations will want to further increase the stock.

Proposition 5.3. *Let $D_1^M = D^N$ given by (5.13) and D_2^M be given by*

$$u'\left(\alpha D_2^M\right) = \frac{\psi\left(D_2^M\right)}{\delta+\alpha}\left(u\left(\alpha D_2^M\right) - \underline{u}\right) \tag{5.14}$$

Define

$$W^M(D) = \int_t^\infty \left(u(z(s))\left[1-(H(s)-H(t))\right] + \underline{u}\left[H(s)-H(t)\right]\right)e^{-\delta(s-t)}ds$$
$$s.t.\ \dot{D} = z-\alpha D,\ D(t)=D_t,\ D(s)\le D^M\ \forall\ s\ge t,\ \dot{H} = \psi(D)(z-\alpha D) \tag{5.15}$$

There exists a continuum of Markov equilibria indexed by $D^M \in \left[D_1^M, D_2^M\right]$ such that

$$\zeta^M(D) = \begin{cases} \text{argmax}_{z(t)}\, W^M(D) & \text{if } D < D^M \\ \alpha D & \text{if } D \ge D^M \end{cases} \tag{5.16}$$

[68] For convenience, I present the analysis of the Markov equilibrium in this section and in section 5.4 in continuous time. The equilibria presented are the limiting cases of the equilibria when each generation is alive for a period of length ϵ and ϵ goes to zero.

When generations have consistent beliefs about the steady-state stock, the beliefs become self-fulfilling, even if they result in an inefficient equilibrium $D^M > D^N$. The upper bound of the equilibrium range D_2^M may either increase or decrease in α, as in the naive solution. As opposed to D_1^M, the upper bound unambiguously increases in δ: D_2^M does not depend on ξ, so the only effect of a higher discount rate is to increase the weight of current utility gains from increasing the stock compared to the stream of possible utility losses $\left(u\left(\alpha D_2^M\right) - \underline{u}\right)/\delta$. The $D^M = D^N$ equilibrium yields the highest welfare for all generations as it comes closest to internalizing the intrinsic welfare loss from a catastrophe. When $D^M = D_2^M$, each regulator behaves as if he does not care about the long-run future ($\xi = 0$). By contrast, in the naive solution each generation believes it decides the steady-state stock. Since it is in no generation's interest to exceed D^N, $D(t) > D^N$ is ruled out.

Corollary 5.2. *The first generation's welfare in the naive solution is lower than in the Markov solution when $D^M = D^N$.*

The naive solution suffers from a different inefficiency. Generation t mistakenly perceives the steady-state stock to be $D^{t,N} < D^N$, so its emissions do not maximize expected discounted utility under the correct belief D^N. In the Markov solution with $D^M = D^N$, all generations have consistent beliefs, so the emissions path does maximize $\int_t^\infty \left(u\left(z\left(s\right)\right)\left[1 - \left(H\left(s\right) - H\left(t\right)\right)\right] + \underline{u}\left[H\left(s\right) - H\left(t\right)\right]\right) e^{-\delta(s-t)} ds$ subject to $D\left(s\right) \leq D^N \ \forall \ s \geq t$. Figure 5.1 illustrates emissions and stocks in the three scenarios. Emissions in the naive solution are initially close to those in the commitment solution, but increasingly diverge as future generations put more weight on their own utility than their predecessors. The Markov solution converges to the same maximum stock as the naive solution, but the maximum is attained much earlier, resulting in higher welfare for early generations than in the naive solution.

Proposition 5.4. *At each point in time, the emission stock is lower in the commitment solution than in the naive and Markov solutions: $D^C\left(t\right) \leq D^N\left(t\right) \leq D^M\left(t\right) \ \forall \ t > 0$.*

Regulators in the commitment and naive solutions maximize a weighted sum of expected utility and catastrophe risk, so the optimal path is the same as in a constrained optimization problem in which the regulator maximizes expected discounted utility subject to the stock not exceeding an exogenous ceiling D^C or $D^{t,N}$ at any point in time (see

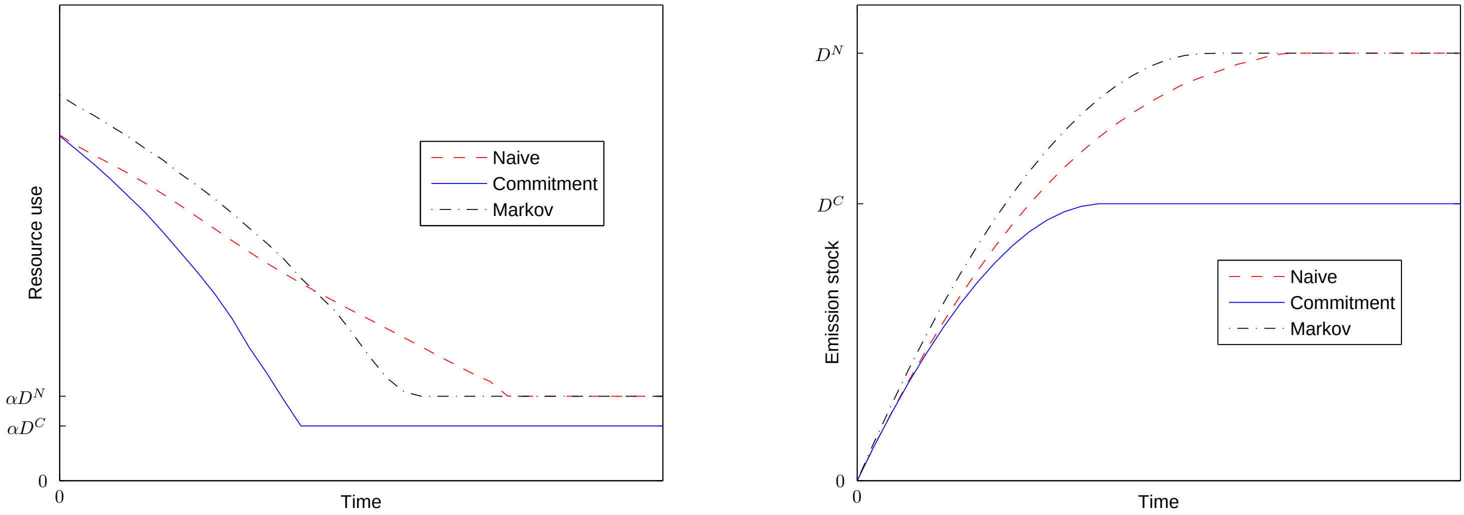

Figure 5.1: Emission flows (left) and stocks (right) in commitment, naive and Markov solutions

Chakravorty et al. (2006, 2008)). By Proposition 5.3, Markovian regulators also solve a constrained optimization problem in equilibrium. The 'carbon budget' is larger in the naive and Markov solutions, so conditional on the stock D the emission flows are higher than in the commitment solution. Because emissions can be ranked for any given stock, the stocks can also be ranked unambiguously at each point in time.

The progress on prominent objectives such as biodiversity preservation and limiting climate change has so far not been encouraging. Current policymakers care less about future consumption than future policymakers do, so the environment is best served when the current generation has full commitment power. In the absence of commitment, a catastrophe becomes more likely because future generations are unwilling to comply with current plans of 'pollute now, clean up later'. Fully rational policies lead to the fastest degradation: because rational decision makers realize that their successors are not more willing to pay for the environment than they are, the long-term objective of limiting catastrophe risk to acceptable levels is out of reach and it is optimal to continue under business-as-usual.

The dismal results in this model rely on a large number of generations having an unlimited ability to pollute. Section 5.2 varied the number of generations; the next section considers pollutant scarcity.

5.4. Infinite horizon, scarce pollutant

In this section, I analyze optimal emissions when the pollutant is scarce. Cumulative emissions (i.e. pollutant consumption) cannot exceed a resource supply S. Unless otherwise noted, I preserve the notation from section 5.3. Let $D_{\max}(t) \equiv \max_{s<t} D(s)$ denote the maximum stock that has been reached until time t and $\tau \equiv \operatorname{argmin}_t \left\{ D_{\max}(t) \geq \hat{D} \right\}$ be the occurrence time of the catastrophe. For simplicity, and because the resource constraint already limits post-catastrophe utility, I abstract from direct utility reductions after a catastrophe. Generation t's welfare is

$$W^t(S(t), D(t), D_{\max}(t)) = \int_t^{\infty} u(z(s)) e^{-\delta(s-t)} ds - \xi \frac{P[\tau \in [t, \infty)]}{1 - P[\tau \in [0, t)]}$$

$$\text{s.t. } \dot{S} = -z, \ \dot{D} = z - \alpha D, \ \dot{D}_{\max} = \mathbf{1}_{\{D = D_{\max}\}} (z - \alpha D), \ S, D, D_{\max} \geq 0 \tag{5.17}$$

When the remaining resource supply is sufficiently small compared to the current pollution stock, the Hotelling extraction path that maximizes discounted utility can be followed without catastrophe risk. Optimal extraction falls quickly enough over time so that the current 'safe' pollution stock is never exceeded. Because catastrophe risk is the only source of time inconsistency, this result applies to the commitment, naive and Markov solutions. I formalize this result in the next Lemma, after introducing some notation. Let

$$\mathcal{B} \equiv \left\{ (S, D) : \underset{z(t)}{\operatorname{argmax}} \left\{ \int_t^{\infty} u(z(s)) e^{-\delta s} ds \text{ s.t. } \dot{S} = -z, \ S \geq 0 \right\} = \alpha D \right\}$$

denote the combinations of S and D for which the emissions $z(t)$ that maximize discounted utility (disregarding catastrophe risk) equal the natural decay of the current stock αD. Define $S_{\mathcal{B}} : \mathbb{R}_+ \to \mathbb{R}_+$ as $\{S : (D, S) \in \mathcal{B}\}$. Given a pollution stock D, $S_{\mathcal{B}}$ is the level of resource supply such that the combination (S, D) is in the set $\mathcal{B}$. $S_{\mathcal{B}}$ is an increasing function: the higher the pollution stock D, the higher the remaining resource supply for which the discounted-utility maximizing $z(t)$ equals αD.

Lemma 5.1. *If generation t inherits $S \leq S_{\mathcal{B}}(D)$, the commitment, naive and Markov* *solutions to* (5.17) *are equal to that of a standard Hotelling problem*

$$\max_z \left\{ W^H(S) = \int_t^{\infty} u(z(s)) e^{-\delta s} ds \ s.t. \ \dot{S} = -z, \ S \geq 0 \right\} \tag{5.18}$$

I assume that the pollution stock is non-decreasing before the terminal phase in which extraction follows a Hotelling path and the catastrophe hazard is zero. The intuition behind this assumption is similar to section 5.3. If it is worthwhile to increase the stock and risk a catastrophe at time t, it can only be optimal to reduce the stock at $t' > t$ if it is necessitated by a dwindling resource supply. Lemma 5.2 shows that the terminal phase is preceded by a non-degenerate interval in which the pollution stock is constant. This result too applies to all three (commitment, naive and the Markov) solutions. The marginal cost of emissions is discontinuous at $z = \alpha D$ when $D = D_{\max}$ in all three solutions. When the system is close to the terminal phase, the benefit of increasing the stock is small. As a result, even far-future generations are hesitant to risk a catastrophe.

Lemma 5.2. *Suppose that $D = D_{max}$ and $S = S_{\mathcal{B}}(D) + \epsilon$, ϵ small. Let $W(S, D)$ be the welfare function when $D = D_{max}$. Then* $\text{argmax}_z W(S, D) = \alpha D$.

Lemmas 5.1 and 5.2, together with the assumption that the stock is non-decreasing before the terminal phase, divide the time horizon into three regimes for all (commitment, naive and Markov) solutions: a first regime with an increasing pollution stock, a second with a constant stock and a third with a declining stock. Lemma 5.2 characterizes the boundary between the second and third regime; I now turn to the boundary between the first and second regime, i.e. the maximum value of S for which the pollution stock is kept constant for a given D. Unlike the minimum value of S for which $z = \alpha D$ for a given D from Lemma 5.2, the maximum is not equal across the commitment, naive and Markov solutions: the shadow cost of pollution plays an important role in the decision when to stabilize the stock, and this cost is higher in the commitment solution than in the naive and Markov solutions. Again, I introduce some auxiliary notation. Define

$\tilde{W}^k(S,D)\,,\; k\in\{\{C,t'\},N,M\}:\{(S,D):S>S_{\mathcal{B}}(D)\}\to\mathbb{R}_+$ as

$$\begin{aligned}
\tilde{W}^{C,t'}(S(t),D(t)) &= \int_t^\infty \left(u(z(s))\left[1-H(s)\right]+\underline{u}H(s)-\eta(s)\,\xi e^{\delta t'}\right)e^{-\delta(s-t)}ds\\
&\quad \text{s.t. } \dot{S}=-z,\; \dot{D}=z-\alpha D,\; \dot{H}=\psi(D)(z-\alpha D)\,,\; S,D\geq 0\\
&\quad z(s')-\alpha D(s')\gtreqless 0\;\forall\; s'\lesseqgtr t'\\
\tilde{W}^{N}(S(t),D(t)) &= \int_t^\infty \left(u(z(s))\left[1-H(s)\right]+\underline{u}H(s)-\eta(s)\,\xi\right)e^{-\delta(s-t)}ds\\
&\quad \text{s.t. } \dot{S}=-z,\; \dot{D}=z-\alpha D,\; \dot{H}=\psi(D)(z-\alpha D)\,,\; S,D\geq 0\\
\tilde{W}^{M}(S,D) &= \int_t^\infty \left(u(z(s))\left[1-H(s)\right]+\underline{u}H(s)-\eta(s)\,\xi\right)e^{-\delta(s-t)}ds\\
&\quad \text{s.t. } \dot{S}=-z,\; \dot{D}=z-\alpha D,\; \dot{H}=\psi(D)(z-\alpha D)\,,\; S,D\geq 0\\
&\quad z(s)=\zeta^M(S(s),D(s))\;\;\forall\; s>t
\end{aligned}$$

If the initial generation commits to stabilizing the stock exactly at time t', the welfare of a fictitious generation $t\leq t'$ that shares the initial generation's preference for catastrophe prevention is equal to $\tilde{W}^{C,t'}$. If generation t is the first generation that keeps the stock constant in the naive or Markov solution, its welfare is equal to $\tilde{W}^N$ or $\tilde{W}^M$, respectively. Similar to the model with an abundant pollutant, the initial generation simultaneously decides on the triplet $(t',S(t'),D(t'))$ at which it will stabilize the stock in the commitment solution, but the combinations (S,D) at which the stock can be stabilized in the naive and Markov solutions do not depend on time. Now I can define the combinations (S,D) that mark the boundary between the values of (S,D) for which the pollution stock increases, and for which it remains constant. Let

$$\mathcal{A}^i\equiv\left\{(S,D):S=\underset{S'}{\operatorname{argmax}}\left\{\underset{z(t)}{\operatorname{argmax}}\,\tilde{W}^k(S',D)=\alpha D\right\}\right\},\; k\in\{\{C,t'\},N,M\}$$

and define $S_{\mathcal{A}^k}:\mathbb{R}_+\to\mathbb{R}_+$ as $\left\{S:(D,S)\in\mathcal{A}^k\right\},\; k\in\{\{C,t'\},N,M\}$ as the value of S for which (S,D) is in $\mathcal{A}^k$ for a given D.

Lemma 5.3. $S_{\mathcal{A}^{C,t'}}(D)>S_{\mathcal{A}^N}(D)\geq S_{\mathcal{A}^M}(D)$

The literal interpretation of Lemma 5.3[69] is of limited direct interest, but the Lemma

[69] If the initial generation were to stabilize the pollution stock at some D under commitment, it will have a larger resource supply remaining when reaching this D than naive or Markovian generations would have if they were to stabilize pollution at the same level of D.

is useful for a graphical intuition of the extraction paths in the commitment, naive and Markov solutions. Figure 5.2 shows the movement through the state space along the optimal path in the commitment solution (the (S, D) combinations that are in the sets $\mathcal{B}$ and $\mathcal{A}^k$ are on increasing but not necessarily straight lines). Starting from the initial condition, the pollution stock increases and the resource supply declines, resulting in a northwest movement in the (S, D) plane until the state reaches a point on the rightmost solid line. From then on, the pollution stock remains constant and the resource supply declines, giving rise to a westward movement until the state is in the set $\mathcal{B}$. In the last phase, the pollution stock and resource supply both decline. In the naive and Markov solutions, the first phase (in which the pollution stock increases) continues until the state reaches a point on the dashed line, which is strictly to the northwest of the line that marks the transition to the second regime in the commitment solution.

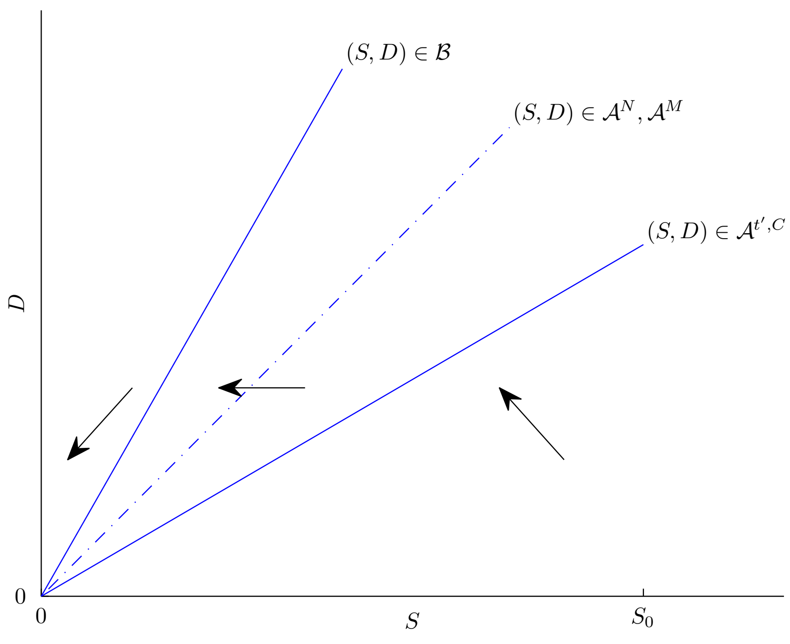

Figure 5.2: Movement through the state space along the optimal path

An analytical comparison of the commitment, naive and Markov paths is beyond the scope of this paper. In section 5.3, the commitment and Markov paths were similar in the sense that they both maximized expected discounted utility subject to the stock remaining below an exogenous ceiling - the only difference being the value of this exogenous ceiling. Also with a scarce pollutant, the commitment path looks like the solution of a time-consistent constrained optimization problem, with the value of the ceiling depending on the initial generation's choice of $(t', S(t'), D(t'))$. The Markov solution will look different however. The intuition is that there is a unique point in $\mathcal{A}^M$ that can be approached from the initial state as the solution to a time-consistent constrained optimization problem.[70] From the initial generation's perspective, the pollution stock at this point is too high.[71] If the initial generation believes that subsequent generations will behave as if they solve a time-consistent constrained optimization problem, it can profitably deviate by decreasing its resource use, which results in a lower maximum pollution stock. Though the resource supply is still exhausted eventually, emissions are spread more evenly over time. This allows the natural decay to reduce the maximum carbon stock, and hence the probability of a catastrophe. Hence, the maximum stock is higher in Markov equilibrium than under commitment, but the maximum is approached in a comparatively slower fashion.

The results in section 5.3 (in which the pollutant is abundant) are a limiting case of the model with a scarce pollutant. When the initial resource supply is sufficiently large, the actions of early generations will be similar to section 5.3. I perform a simulation to illustrate emissions in the commitment and Markov solutions when the resource supply is limited. I use a quadratic utility function and a discrete grid for $(S, D, D_{\max})$. Figure 5.3 depicts the results. In this example, initial emissions are lower in Markov equilibrium than under commitment, because of the first generation's incentive to reduce emissions outlined at the end of the previous paragraph.

5.5. Conclusion

It is well known that discounted utilitarianism can recommend environmental degradation as optimal policy. This paper shows that welfare criteria that explicitly value the long-run future may also not prevent a catastrophe when the environmental problem is long-lived and caused by abundant pollutants. Future generations will not reduce their

[70] The (S, D) combinations in $\mathcal{A}^k$ are positively correlated, whereas the (S, D) combinations such that the pollution stock reaches the exogenous ceiling D in a time-consistent constrained optimization problem when the remaining resource supply equals S are negatively correlated.

[71] The pollution stock is higher than the level at which the stock is stabilized under commitment, and the remaining resource supply at the moment of stabilization is lower than under commitment.

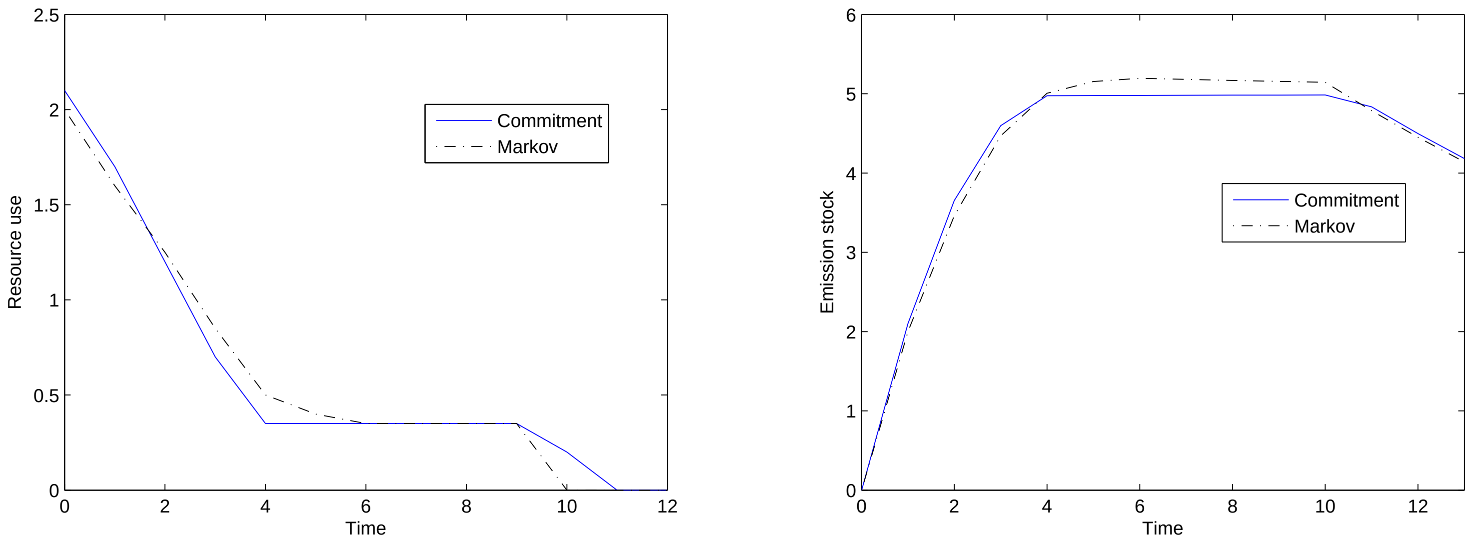

Figure 5.3: Emission flows (left) and stocks (right) in commitment and Markov solutions

consumption to stabilize pollution concentrations at the current generation's preferred level. As a result, rational policymakers conclude that mitigation is futile, and equilibrium behaviour may look as if each policymaker has no intrinsic desire for catastrophe prevention. Given the large reserves of coal and unconventional oil, this is a worrying message for limiting climate change. My results suggest that if today's generation wants to enact its preferences and if commitments through policy rules are not possible, its best chance is to develop a technological commitment device such as a substitute for these abundant fossil fuels - rather than reducing consumption and hoping that future generations will do the same.

The paper also suggests that instrumental and intrinsic catastrophe aversion have different implications for equilibrium policies. Generations are more likely to preserve the environment if they value its contribution to their descendants' utility, for example because ecosystem services are valuable in production and consumption, than if they care for the environment for its own sake. Environmental amenities that have no economic value are more likely to be sacrificed by future generations that care more about their own consumption, which in turn makes preservation by the current generation less worthwhile.

5.A. Appendix

5.A.1. Two-period model: ranking z_1^C and z_1^M

Lemma 5.4 provides unambiguous rankings of z_1^C and z_1^M for iso-elastic and quadratic utility when the catastrophe threshold follows a uniform distribution. When $\hat{D}$ is uniformly distributed, the terms III in the first-order conditions (5.6) and (5.2a) are equal. We are thus left with the terms II, which make the first generation more conservationist in Markov equilibrium, and the strategic term IV. The sufficient condition $f'(z_1 + r(z_1)) \geq 0$ for the second generation's reaction function to be downward sloping is satisfied for a uniformly distributed catastrophe threshold. The strategic effect thus encourages the first generation to emit more. This typically raises cumulative emissions,[72] but changes the ratio between first- and second-period marginal utilities to the first generation's benefit. For iso-elastic utility, the strategic effect dominates the effect from term II, and emissions are higher in Markov equilibrium than under commitment. For quadratic utility, the converse applies and the first generation is more prudent in the Markov solution. The intuition is that with quadratic utility, the second generation's utility is more concave in prices than the quantity demanded (i.e. the reaction function) is, compared to under iso-elastic utility. Increasing first-period emissions, which raises the effective price of second-period consumption, therefore strongly affects the second generation's utility but not so much the quantity demanded in case of quadratic utility. This makes it less attractive to increase first-period consumption in Markov equilibrium than in the case of iso-elastic utility.

Lemma 5.4. *Let $\hat{D}$ be uniformly distributed ($F(D) = D/\bar{D}$ and $f(D) = 1/\bar{D}$) and ξ sufficiently large so that $z_1^C < \infty$, $z_1^M < \infty$. For iso-elastic utility $u_t(z_t) = \frac{z_t^{1-\eta}}{1-\eta}$, $z_1^M > z_1^C$. For quadratic utility $u_t(z_t) = az_t - \frac{1}{2}bz_t^2$, $z_1^M < z_1^C$.*

Proof. First, consider iso-elastic utility $u_t(z_t) = \frac{z_t^{1-\eta}}{1-\eta}$. If the catastrophe has not occurred by the start of the second period, we have

$$w_2^M = \frac{\left(z_2^M\right)^{1-\eta}}{1-\eta} - \xi \, \frac{z_1^M + z_2^M}{\bar{D} - z_1^M} \Leftrightarrow \left(z_2^M\right)^{-\eta} = \frac{\xi}{\bar{D} - z_1^M} \Leftrightarrow r(z_1) = \left(\frac{\bar{D} - z_1}{\xi}\right)^{\frac{1}{\eta}}$$

[72] The proof contains functional forms for the reaction functions, from which one can derive conditions for $|r'(z_1)| < 1$

Substituting in (5.1a), I obtain

$$w_1^M = \frac{\left(z_1^M\right)^{1-\eta}}{1-\eta} + \rho \frac{\left(\frac{\bar{D}-z_1}{\xi}\right)^{\frac{1-\eta}{\eta}}}{1-\eta}\left(1-\frac{z_1^M}{\bar{D}}\right) + \frac{z_1^M \rho \underline{u}}{\bar{D}} - \xi \frac{z_1^M + \left(\frac{\bar{D}-z_1^M}{\xi}\right)^{\frac{1}{\eta}}}{\bar{D}}$$

The associated first-order condition is

$$\left(z_1^M\right)^{-\eta} - \frac{\rho}{\bar{D}}\left(\frac{\left(\frac{\bar{D}-z_1^M}{\xi}\right)^{\frac{1-\eta}{\eta}}}{1-\eta} - \underline{u}\right) + \frac{1-\rho}{\eta\bar{D}}\left(\frac{\bar{D}-z_1^M}{\xi}\right)^{\frac{1-\eta}{\eta}} - \frac{\xi}{\bar{D}} = 0 \tag{5.19}$$

Conversely, in the commitment outcome second-period emissions satisfy

$$w_2^C = \rho\frac{\left(z_2^C\right)^{1-\eta}}{1-\eta} - \xi\frac{z_1^C + z_2^C}{\bar{D}-z_1^C} \Leftrightarrow \rho\left(z_2^C\right)^{-\eta} = \frac{\xi}{\bar{D}-z_1^C} \Leftrightarrow z_2^C = \left(\frac{\rho\left(\bar{D}-z_1^C\right)}{\xi}\right)^{\frac{1}{\eta}}$$

This gives us

$$w_1^C = \frac{\left(z_1^C\right)^{1-\eta}}{1-\eta} + \rho\frac{\left(\frac{\rho\left(\bar{D}-z_1^C\right)}{\xi}\right)^{\frac{1-\eta}{\eta}}}{1-\eta}\left(1-\frac{z_1^C}{\bar{D}}\right) + \frac{z_1^C\rho\underline{u}}{\bar{D}} - \xi\frac{z_1^C + \left(\frac{\rho\left(\bar{D}-z_1^C\right)}{\xi}\right)^{\frac{1}{\eta}}}{\bar{D}}$$

and FOC

$$\left(z_1^C\right)^{-\eta} - \frac{\rho}{\bar{D}}\left(\frac{\left(\frac{\rho\left(\bar{D}-z_1^C\right)}{\xi}\right)^{\frac{1-\eta}{\eta}}}{1-\eta} - \underline{u}\right) - \frac{\xi}{\bar{D}} = 0 \tag{5.20}$$

It can be shown that the left-hand side of (5.19) is larger than the left-hand side of (5.20) for all z_1 and $\rho \in (0,1)$. Therefore, $z_1^M > z_1^C$.

Now consider quadratic utility $u_t(z_t) = az_t - \frac{1}{2}bz_t^2$. If the catastrophe has not occurred by the start of the second period, the second generation's welfare is

$$w_2^M = az_2^M - \frac{1}{2}b\left(z_2^M\right)^2 - \xi\frac{\frac{z_2^M}{\bar{D}}}{1-\frac{z_1^M}{\bar{D}}} \Leftrightarrow a - bz_2^M - \frac{\xi}{\bar{D}\left(1-\frac{z_1^M}{\bar{D}}\right)} = 0 \Leftrightarrow r(z_1) = \frac{a\left(\bar{D}-z_1\right)-\xi}{b\left(\bar{D}-z_1\right)}$$

Substituting in (5.1a), I obtain

$$w_1^M = az_1^M - \frac{1}{2}b\left(z_1^M\right)^2 + \rho\left(a\left(\frac{a\left(\bar{D}-z_1\right)-\xi}{b\left(\bar{D}-z_1\right)}\right) - \frac{1}{2}b\left(\frac{a\left(\bar{D}-z_1\right)-\xi}{b\left(\bar{D}-z_1\right)}\right)^2\right)\left(1-\frac{z_1^M}{\bar{D}}\right)$$
$$+\frac{z_1^M\rho\underline{u}}{\bar{D}} - \xi\,\frac{z_1+\frac{a\left(\bar{D}-z_1\right)-\xi}{b\left(\bar{D}-z_1\right)}}{\bar{D}}$$

The first-order condition is

$$a - bz_1^M - \frac{1}{2}\rho\frac{\left(a^2-2b\underline{u}\right)\left(\bar{D}-z_1^M\right)^2-\xi^2}{b\bar{D}\left(\bar{D}-z_1^M\right)^2} + \frac{\xi^2\left(1-\rho\right)}{b\bar{D}\left(\bar{D}-z_1^M\right)^2} - \frac{\xi}{\bar{D}} = 0$$

In the commitment outcome, the first generation chooses z_2^C to maximize

$$w_2^C = \rho\left(az_2^C - \frac{1}{2}b\left(z_2^C\right)^2\right) - \xi\,\frac{\frac{z_2^C}{\bar{D}}}{1-\frac{z_1^C}{\bar{D}}} \Leftrightarrow \rho\left(a-bz_2^C\right) - \frac{\xi}{\bar{D}\left(1-\frac{z_1^C}{\bar{D}}\right)} = 0 \Leftrightarrow z_2^C = \frac{\rho a\left(\bar{D}-z_1^C\right)-\xi}{\rho b\left(\bar{D}-z_1^C\right)}$$

The first generation's welfare is then

$$w_1^C = az_1^C - \frac{1}{2}b\left(z_1^C\right)^2 + \rho\left(a\left(\frac{\rho a\left(\bar{D}-z_1^C\right)-\xi}{\rho b\left(\bar{D}-z_1^C\right)}\right) - \frac{1}{2}b\left(\frac{\rho a\left(\bar{D}-z_1^C\right)-\xi}{\rho b\left(\bar{D}-z_1^C\right)}\right)^2\right)\left(1-\frac{z_1^C}{\bar{D}}\right)$$
$$+\frac{z_1^C\rho\underline{u}}{\bar{D}} - \xi\,\frac{z_1^C+\frac{\rho a\left(\bar{D}-z_1^C\right)-\xi}{\rho b\left(\bar{D}-z_1^C\right)}}{\bar{D}}$$

giving rise to the following first-order condition

$$a - bz_1^C - \frac{1}{2}\frac{\left(\bar{D}-z_1^C\right)^2\left(a^2-2b\underline{u}\right)\rho^2-\xi^2}{\rho b\bar{D}\left(\bar{D}-z_1^C\right)^2} = 0 \tag{5.21}$$

Letting $z_1^C = z_1^M = z_1$, we have

$$\frac{\partial w_1^C}{\partial z_1} - \frac{\partial w_1^M}{\partial z_1} = \frac{1}{2}\frac{\xi^2\left(1-\rho\right)^2}{\rho b\bar{D}\left(\bar{D}-z_1\right)^2} > 0$$

Therefore, for quadratic utility $z_1^C > z_1^M$. <u>Q.E.D.</u>

5.A.2. Proofs

Proof of Proposition 5.1

Proof. I omit the superscript C except to indicate the steady state. From (5.10) it is apparent that if $z(t) = \alpha D(t)$ for some t, we must also have $z(s) = \alpha D(s)\ \forall\ s > t$. Otherwise, the first generation could improve its welfare by choosing $z(t) > \alpha D(t)$, as the current value cost of triggering a catastrophe is lower at t than at s. Moreover, the pollution stock must stabilize at some finite level because $\lim_{z\to\infty} u'(z) = 0$, $\lim_{D\to\infty} \psi(D) >> 0$ and since $D(s)$ is monotonic along the optimal path. Combining the above observations, there exists some t' such that $D(t') = D^C$ and $z(t) = \alpha D^C\ \forall\ t \geq t'$.

Now consider the alternative problem

$$
\begin{aligned}
\max_z \Bigg\{ \tilde{W}^C(D(t)) = &\int_t^\infty \left(u(z(s))\left[1 - (H(s) - H(t))\right] + \underline{u}\left[H(s) - H(t)\right] - \eta(s)\,\xi e^{\delta t'} \right) e^{-\delta s} ds \\
&s.t.\ \dot{D} = z - \alpha D,\ D(t) = D_t,\ \dot{H} = \psi(D)(z - \alpha D) \\
&= \int_t^\infty \left(u(z(s))\left[1 - (H(s) - H(t))\right] + \left(\underline{u} - \delta\xi e^{\delta t'}\right)\left[H(s) - H(t)\right] \right) e^{-\delta s} ds \\
&s.t.\ \dot{D} = z - \alpha D,\ D(t) = D_t,\ \dot{H} = \psi(D)(z - \alpha D) \Bigg\}
\end{aligned} \tag{5.22}
$$

In the above problem, the current-value welfare cost of triggering a catastrophe at time s is always equal to $\xi e^{\delta t'}$, whereas this cost depends on the time of occurrence $(\xi e^{\delta s})$ in the original problem (5.10). The above problem has the same solution as (5.10) evaluated at $D(t) = D(t')$, but (5.22) is stationary whereas (5.10) is not. The derivatives of $W^C(D(t'))$ and $\tilde{W}^C(D(t'))$ with respect to $z(t')$ have the same sign. Because (5.22) is stationary, I can analyze its steady state, assuming it is approached by a path in which $D(s)$ is non-decreasing. t' and $D(t') = D^C$ satisfy the conditions in the proposition text if and only if $z = \alpha D^C$ is the optimal steady-state policy in (5.22). Let $\tilde{v}$ and $\tilde{\mu}$ denote the costate variables for $\tilde{V}(D(t)) = \max_z \tilde{W}(D(t))$ and D, respectively. From

Appendix 5.A.3, the steady-state conditions are

$$\dot{D} = z - \alpha D = 0 \tag{5.23a}$$

$$\dot{\tilde{\mu}} = (\delta + a)\tilde{\mu} + \psi(D)(z - \alpha D)\tilde{\mu} - \psi(D)\alpha\left(\tilde{v} - \frac{\underline{u}}{\delta} + \xi e^{\delta t'}\right) = 0 \tag{5.23b}$$

$$\dot{\tilde{v}} = \delta\tilde{v} - u(z) + \psi(D)(z - \alpha D)\left(\tilde{v} - \frac{\underline{u}}{\delta} + \xi e^{\delta t'}\right) = 0 \tag{5.23c}$$

$$u'(z) + \tilde{\mu} - \psi(D)\left(\tilde{v} - \frac{\underline{u}}{\delta} + \xi e^{\delta t'}\right) = 0 \tag{5.23d}$$

Solving (5.23) for D, $\tilde{\mu}$, $\tilde{v}$ and z yields

$$u'(\alpha D) = \frac{\psi(D)}{\delta + \alpha}\left[u(\alpha D) - \underline{u} + \delta\xi e^{\delta t'}\right]$$

Therefore, t' and D^C must satisfy (5.11). Q.E.D.

Proof of Proposition 5.2

Proof. By the argument in the main text, the steady-state stock cannot exceed D^N. Consider a generation t that inherits stock $D(t) < D^N$. Let $D^{t,N}(t')$ and $z^{t,N}(t')$ denote the stock and emissions respectively at time $t' > t$ in generation t's preferred path. Suppose that $D^{t,N} = D^N$ and $D^{t,N}(t') = D^{t,N}$.[73] Analogous to the proof of Proposition 5.1, it can only be optimal to choose $z^{t,N}(t') = \alpha D^N$ iff

$$u'\left(\alpha D^N\right) = \frac{\psi\left(D^N\right)}{\delta + \alpha}\left[u\left(\alpha D^N\right) - \underline{u} + \delta\xi e^{\delta(t'-t)}\right] \tag{5.24}$$

If (5.13) holds at D^N, the right hand side of (5.24) exceeds the left hand side at $D^{t,N} = D^N$ since $t' > t$. By Assumptions 5.1 and 5.2, we must therefore have $D^{t,N} < D^N$.

I complete the proof of $\lim_{t\to\infty} D^{t,N} = D^N$ by noting that whenever $D^{t,N} < D^N$ and $D^{t,N}(t') = D^{t,N}$, generation $t' > t$ prefers $D^{t',N} > D^{t,N}$. $D^{t,N}(t') = D^{t,N}$ implies

$$u'\left(\alpha D^{t,N}\right) = \frac{\psi\left(D^{t,N}\right)}{\delta + \alpha}\left[u\left(\alpha D^{t,N}\right) - \underline{u} + \delta\xi e^{\delta(t'-t)}\right] \tag{5.25}$$

[73] If $D^{t,N} = D^N$ but $D(t)$ does not reach D^N in finite time, a modified version of the below argument still applies: for t' arbitrarily large and ϵ arbitrarily small, the left hand side of (5.24) evaluated at $D^N - \epsilon$ is larger than the right hand side.

If $D^{t',N} = D^{t,N}$, we must also have

$$u'\left(\alpha D^{t,N}\right) = \frac{\psi\left(D^{t,N}\right)}{\delta+\alpha}\left[u\left(\alpha D^{t,N}\right) - \underline{u} + \delta\xi\right] \tag{5.26}$$

Clearly, (5.25) and (5.26) cannot hold simultaneously. When (5.25) holds, the left hand side of (5.26) is larger than the right hand side at $D^{t,N}$: generation t cares less about consumption at $t' > t$ than generation t' does, so they cannot both intend to stabilize at $D^{t,N}$ at time t'. Generation t' will choose $D^{t',N} > D^{t,N}$, so $z^{t',N}(t') > \alpha D^{t,N}$. As the stock approaches D^N, the target levels $D^{t,N}$ must also approach D^N. The comparative statics in the proposition texts follow by total differentiation. For α,

$$\begin{aligned}\left(\alpha\frac{\partial D^N}{\partial\alpha}\right)u''\left(\alpha D^N\right) =& \frac{\frac{\partial D^N}{\partial\alpha}\psi'\left(D^N\right)(\alpha+\delta) - \psi\left(D^N\right)}{(\delta+\alpha)^2}\left[u\left(\alpha D^N\right) - \underline{u} + \delta\xi\right] \\ &+ \frac{\psi\left(D^N\right)}{\delta+\alpha}\left[\alpha\frac{\partial D^N}{\partial\alpha}u'\left(\alpha D^N\right)\right]\end{aligned}$$

After some rearranging, I obtain

$$\frac{\partial D^N}{\partial\alpha} = \frac{-(\alpha+\delta)^2 D^N u''\left(\alpha D^N\right) + (\alpha+\delta)D^N\psi\left(D^N\right)u'\left(\alpha D^N\right) - \psi\left(D^N\right)\left(u\left(\alpha D^N\right) - \underline{u} + \delta\xi\right)}{(\alpha+\delta)\left[\alpha(\alpha+\delta)u''(\alpha D^N) - \alpha\psi(D^N)u'(\alpha D^N) - \psi'(D^N)(u(\alpha D^N) - \underline{u} + \delta\xi)\right]} \tag{5.27}$$

All terms in the denominator are negative by Assumptions 5.1 and 5.2. The total effect depends on the sign of the numerator. For δ,

$$\begin{aligned}\alpha\frac{\partial D^N}{\partial\delta}u''\left(\alpha D^N\right) =& \frac{\frac{\partial D^N}{\partial\delta}\psi'\left(D^N\right)(\alpha+\delta) - \psi\left(D^N\right)}{(\delta+\alpha)^2}\left[u\left(\alpha D^N\right) - \underline{u} + \delta\xi\right] \\ &+ \frac{\psi\left(D^N\right)}{\delta+\alpha}\left[\alpha\frac{\partial D^N}{\partial\delta}u'\left(\alpha D^N\right) + \xi\right]\end{aligned}$$

Rearranging gives

$$\frac{\partial D^N}{\partial\delta} = \frac{\psi\left(D^N\right)\left[-u\left(\alpha D^N\right) + \underline{u} + \alpha\xi\right]}{(\alpha+\delta)\left[\alpha(\alpha+\delta)u''(\alpha D^N) - \alpha\psi(D^N)u'(\alpha D^N) - \psi'(D^N)(u(\alpha D^N) - \underline{u} + \delta\xi)\right]} \tag{5.28}$$

The denominator is the same as in (5.27). The sign of $\frac{\partial D^N}{\partial\delta}$ is thus determined by the sign of the numerator. <u>Q.E.D.</u>

Proof of Corollary 5.1

Suppose that D^C is reached for the first time at time $t' > 0$. Then it can only be optimal to choose $z^C(t') = \alpha D^C$ iff

$$u'\left(\alpha D^C\right) = \frac{\psi\left(D^C\right)}{\delta+\alpha}\left[u\left(\alpha D^C\right) - \underline{u} + \delta\xi e^{\delta t'}\right]$$

If (5.13) holds at D^N, the right hand side of the above equation exceeds the left hand side at $D^C = D^N$. By Assumptions 5.1 and 5.2, we must therefore have $D^C < D^N$.

Proof of Proposition 5.3

Proof. Recall that D_1^M and D_2^M are unique by Proposition 5.2. I verify that the equilibria in the proposition text satisfy the equilibrium conditions. Let t be sufficiently large and suppose that generation t believes that future generations will follow (5.16) and $D \geq D^M$. Then generation t believes that if it increases the stock, future generations will keep the stock constant.

First, consider the case in which $D^M < D^N$. By Proposition 5.2, generation t would prefer to reach a higher steady-state stock in the naive solution, that is if it could commit all emissions from t onward. I show that this implies that in the Markov solution, generation t will choose $z > \alpha D$. When t is sufficiently large, $D^{t,N}$ is arbitrarily close to D^N. Furthermore, in generation t's preferred path $z^{t,N}(s)$, $D^{t,N}$ is reached in finite time. This means there is exists a $t' > t$ such that

$$(1-\alpha)\, D^{t,N}(t') + z^{t,N}(t') = D^{t,N}$$

and

$$\left.\frac{\partial W^{t,N}\left(D^{t,N}(t')\right)}{\partial z^{t,N}(t')}\right|_{z^{t,N}(t') = D^{t,N} - (1-\alpha) D^{t,N}(t')} = 0 \tag{5.29}$$

The interpretation of (5.29) is that, at $t' > t$ and $D^{t,N}(t') > D(t)$, generation t would choose to increase the pollution stock by $D^{t,N} - (1-\alpha)\, D^{t,N}(t') > 0$ if the stock would remain constant in all subsequent periods. But then by Assumption 5.1 and Assumption 5.2, it must be welfare-improving to increase the stock by the same amount at $D(t)$, given that future generations keep the stock constant at the new level: the marginal utility of consumption is higher, the hazard rate is lower and the current-value cost of a catastrophe is lower. Therefore, $D^M < D^N$ cannot be an equilibrium.

Now turn to the decisions of early generations that inherit a stock $D(t) < D^M$. If generation t believes that subsequent generations will follow (5.16), it realizes that its actions will not affect the maximum stock D^M. When all future generations also believe the maximum stock equals D^M, the preferences of all generations that inherit $D(t) < D^M$ are no longer time-inconsistent. Then the problem of generation t reduces to maximizing the integral of expected discounted utility subject to $D(s) \leq D^M$, i.e.

$$\max_z \int_t^\infty \left(u(z(s))\left[1-(H(s)-H(t))\right]+\underline{u}\left[H(s)-H(t)\right]\right)e^{-\delta s}ds$$
$$s.t.\ \dot{D}=z-\alpha D,\ D(t)=D_t,\ \dot{H}=\psi(D)(z-\alpha D),\ D(s)\leq D^M\ \forall\ s \tag{5.30}$$

The solution to this optimal control problem coincides with the Markov solution. Analogous to Proposition 5.1, the steady state of the unconstrained version of (5.30) - the program that maximizes discounted utility when $\xi = 0$ - satisfies

$$u'(\alpha D)=\frac{\psi(D)}{\delta+\alpha}\left(u(\alpha D)-\underline{u}\right)$$

Therefore, stocks larger than D_2^M are never visited in equilibrium. Q.E.D.

Proof of Proposition 5.4

Proof. Using the results from Propositions 5.1, 5.2 and 5.3, I can rewrite (5.10), (5.12) and (5.15) as constrained optimization problems

$$\max_z \left\{ W^k(D(t)) = \int_t^\infty \left(u(z(s))\left[1-(H(s)-H(t))\right]+\underline{u}\left[H(s)-H(t)\right]\right)e^{-\delta s}ds \right.$$
$$\left. s.t.\ \dot{D}=z-\alpha D,\ D(t)=D_t,\ \dot{H}=\psi(D)(z-\alpha D),\ D(s)\leq D^k\ \forall\ s \right\},\ k\in\{C,\{t,N\},M\} \tag{5.31}$$

where $D^C < D^{t,N} < D^M$ for $0 < t < \infty$. When deciding D^C and $D^{t,N}$ in the commitment and naive solutions, generations trade off discounted utility and the present-value welfare cost of triggering a catastrophe: a higher steady state stock increases the former, but also the expected welfare loss from a catastrophe. Given their choice of D^C or $D^{t,N}$ and the time at which the steady state will be reached, the steady state is approached as in a time-consistent constrained maximization problem in the commitment and naive solutions (as in the Markov solution). I can represent the optimal strategy in each solution as $z = \zeta^k(D) = \zeta\left(D; D^k\right),\ k \in \{C, \{t, N\}, M\}$, where $\zeta^C(D)$ and $\zeta^{t,N}(D)$ are only optimal along the equilibrium path. $D^C < D^{t,N} < D^M$ implies $\zeta^C(D) < \zeta^{t,N}(D) < \zeta^M(D)$ if and only if $\frac{\partial\zeta\left(D;D^k\right)}{\partial D^k} > 0$. Let $V\left(D; D^k\right) \equiv \max_z W^k(D)$ be the value of continuing optimally from stock D subject to $D(s) \leq D^k\ \forall\ s$. Writing $V = V\left(D; D^k\right)$, the HJB equation and the first order condition from the Hamiltonian stipulate

$$\delta V = \max_z \left\{ u(z) + V_D(z - \alpha D) - \psi(D)(z - \alpha D)\left(V - \frac{\underline{u}}{\delta}\right)\right\} \tag{5.32}$$

$$u'(z) + V_D - \psi(D)\left(V - \frac{\underline{u}}{\delta}\right) = 0 \tag{5.33}$$

By (5.32), along the optimal path

$$V_D = \frac{\delta V - u(z)}{z - \alpha D} + \psi(D)\left(V - \frac{\underline{u}}{\delta}\right) \tag{5.34}$$

Substituting (5.34) in (5.33), I obtain

$$\begin{aligned} & u'(z) + \frac{\delta V - u(z)}{z - \alpha D} = 0 \\ & \Leftrightarrow (z - \alpha D)\, u'(z) + \delta V - u(z) = 0 \\ & \Leftrightarrow \tilde{z} u'(\tilde{z} + \alpha D) + \delta V - u(\tilde{z} + \alpha D) = 0 \end{aligned}$$

where $\tilde{z} = z - \alpha D$. Totally differentiate with respect to D^k

$$\begin{aligned} & \frac{\partial \tilde{z}}{\partial D^k} u'(\tilde{z} + \alpha D) + \tilde{z}\frac{\partial \tilde{z}}{D^k} u''(\tilde{z} + \alpha D) + \delta\frac{\partial V}{\partial D^k} - \frac{\partial \tilde{z}}{\partial D^k} u'(\tilde{z} + \alpha D) = 0 \\ & \Leftrightarrow \frac{\partial \tilde{z}}{\partial D^k} \underbrace{\tilde{z} u''(\tilde{z} + \alpha D)}_{<0} + \underbrace{\delta\frac{\partial V}{\partial D^k}}_{>0\ \forall\ D^k < D_2^M} = 0 \end{aligned}$$

By the above, we must have $\frac{\partial \tilde{z}}{\partial D^k} > 0$. Having established $\zeta^C(D) < \zeta^{t,N}(D) < \zeta^M(D)\ \forall\ D$, it automatically follows that $D^C(t) < D^N(t) < D^M(t)$. Q.E.D.

Proof of Lemma 5.1

Proof. I focus on the case $S = S_{\mathcal{B}}(D)$; the proof for $S < S_{\mathcal{B}}(D)$ is analogous. It is sufficient to show that for $z(s) = \text{argmax}_z \left\{ \int_t^\infty u(z(s)) e^{-\delta s} ds \text{ s.t. } \dot{S} = -z,\ S \geq 0 \right\}$, $P\left[\tau < \infty | \hat{D} \leq D(t)\right] = 0$. Suppose $z(s) < \alpha D(s)$ for some $s \geq t$. Then by continuity of D, S and z, there exists a neighborhood (s, s') such that $z(\sigma) \leq \alpha D(\sigma)\ \forall \sigma \in (s, s')$. Conversely, when $z(s) = \alpha D(s)$, there exists a neighborhood (s, s'') such that $z(\sigma) < \alpha D(\sigma)\ \forall\ \sigma \in (s, s'')$ since $\dot{z} < 0$ in the solution to (5.18). Combining these two observations, $z \leq \alpha D$ throughout. Then $D(s) \leq D(t)\ \forall\ s \geq t$, so $P\left[\tau < \infty | \hat{D} \leq D(t)\right] = 0$. Q.E.D.

Proof of Lemma 5.2

Proof. Denote $V(S, D)$ as $\max_z W(S, D)$. When $D = D_{\max}$, the marginal cost of resource consumption is at least $V_S - V_D + \xi\psi(D)$ for $z > \alpha D$. I guess and verify that $\frac{\partial V_S}{\partial S}\big|_{(D,S)\in\mathcal{B}} < 0$. Since $V_D = 0$ and $u'(\alpha D) = V_S$ at $(S, D) = (S_{\mathcal{B}}(D_{\max}), D_{\max})$ by Lemma 5.1, continuity of V_D in S prescribes

$$u'(\alpha D) > V_S$$
$$u'(\alpha D) < V_S - V_D + \xi\psi(D)$$

for S in a neighborhood to the right of $S_{\mathcal{B}}(D)$. This implies $z = \alpha D$. But then there indeed exists a $\frac{\partial V_S}{\partial S}\big|_{(D,S)\in\mathcal{B}} < 0$ such that $z = \alpha D$ satisfies the first order conditions for $S \in (S_{\mathcal{B}}(D), S_{\mathcal{B}}(D) + \epsilon)$ and the Hotelling path is optimal for $S \leq S_{\mathcal{B}}(D)$. Q.E.D.

Proof of Lemma 5.3

Proof. Let $\tilde{V}^k(S, D) \equiv \max_z \tilde{W}^k(S, D),\ k \in \{\{C, t'\}, N, M\}$ and

$$\bar{V}\left(S, \bar{D}\right) = \max_z \left\{ \int_0^\infty u(z(s)) e^{-\delta s} ds \text{ s.t. } \dot{D} = z - \alpha D,\ D(0) = \bar{D},\ D \leq \bar{D}, \dot{S} = -z \right\} \tag{5.35}$$

be the maximum value of discounted utility disregarding catastrophe risk, subject to the constraint that the pollution stock never exceeds the current level.[74] Without loss, let

[74]The characteristics of this problem are discussed in Chakravorty et al. (2006, 2008).

t' be the first moment at which the pollution stock is kept constant (t' may be different between the commitment, naive and Markov solutions). Suppose there exists a (S^*, D^*) such that $(S^*, D^*) \in \mathcal{A}^k,\ k \in \{\{C, t'\}, N, M\}$. Since the regulator in charge of emissions at t' (the initial generation in the commitment solution, and generation t' in the naive and Markov solutions) is indifferent whether or not to increase the stock, we must have

$$\begin{aligned}
u'(\alpha D^*) &= \tilde{V}_S^{C,t'}(S^*, D^*) - \tilde{V}_D^{C,t'}(S^*, D^*) + \psi(D^*)\left(\xi e^{\delta t'} + \tilde{V}^{C,t'}(S^*, D^*) - V^H(S^*)\right) \\
u'(\alpha D^*) &= \tilde{V}_S^{N}(S^*, D^*) - \tilde{V}_D^{N}(S^*, D^*) + \psi(D^*)\left(\xi + \tilde{V}^{N}(S^*, D^*) - V^H(S^*)\right) \\
u'(\alpha D^*) &= \tilde{V}_S^{M}(S^*, D^*) - \tilde{V}_D^{M}(S^*, D^*) + \psi(D^*)\left(\xi + \tilde{V}^{M}(S^*, D^*) - V^H(S^*)\right) \quad (5.36)
\end{aligned}$$

In the commitment and naive solutions, and when $\zeta_S^M(S, D) \leq 0$ in the Markov solution, the regulator in charge at t' knows that future regulators will not further increase the stock. Therefore, the catastrophe hazard is zero in all future periods, so

$$\tilde{V}^{C,t'}(S^*, D^*) = \tilde{V}^N(S^*, D^*) = \tilde{V}^M(S^*, D^*) = \bar{V}(S^*, D^*) \quad (5.37)$$

The marginal value of the resource is equal to that in a setting without catastrophe risk in which the pollution stock is constrained below the current level:

$$\tilde{V}_S^k(S, D)\,|_{(S,D)\in\mathcal{A}^k} = \bar{V}_S(S, D),\ k \in \{\{C, t'\}, N, M\} \quad (5.38)$$

Similarly, the value of increasing the stock by one unit without causing a catastrophe ($\tilde{V}_D^k,\ k \in \{\{C, t'\}, N, M\}$) equals the increase in discounted utility from marginally increasing the exogenous ceiling in (5.35):

$$\tilde{V}_D^k(S, D)\,|_{(S,D)\in\mathcal{A}^k} = \bar{V}_D(S, D),\ k \in \{\{C, t'\}, N, M\} \quad (5.39)$$

By (5.37), (5.38) and (5.39), when the second and third equations in (5.36) hold, we have

$$u'(\alpha D^*) < \tilde{V}_S^{C,t'}(S^*, D^*) - \tilde{V}_D^{C,t'}(S^*, D^*) + \psi(D^*)\left(\xi e^{\delta t'} + \tilde{V}^{C,t'} - V^H(S^*)\right)$$

Hence, there cannot exist a (S^*, D^*) such that $(S^*, D^*) \in \mathcal{A}^k,\ k \in \{\{C, t'\}, N, M\}$. Because $\tilde{V}_{SS}^{C,t'} < 0$, $\tilde{V}_{DS}^{C,t'} > 0$ and $\tilde{V}_S^{C,t'} - V_S^H < 0$, there exists a $S^{**} > S^*$ such that

$$u'(\alpha D^*) = \tilde{V}_S^{C,t'}(S^{**}, D^*) - \tilde{V}_D^{C,t'}(S^{**}, D^*) + \psi(D^*)\left(\xi e^{\delta t'} + \tilde{V}^{C,t'}(S^{**}, D^*) - V^H(S^{**})\right)$$

and hence $(S^{**}, D^*) \in \mathcal{A}^{t',C}$. This establishes $S_{\mathcal{A}^{C,t'}}(D) > S_{\mathcal{A}^N}(D)$. $S_{\mathcal{A}^M}(D) = S_{\mathcal{A}^N}(D)$ fulfills the condition of an equilibrium: in Markov equilibrium, generation t' will not increase the stock if it would not increase the stock in its first-best and if it expects future generations also not to increase the stock. However, if it does expect future generations to increase the stock, it may be optimal to choose $z > \alpha D$, so that $S_{\mathcal{A}^N}(D) > S_{\mathcal{A}^M}(D)$.

Q.E.D.

5.A.3. Piecewise deterministic optimal control

Consider a random variable ε with probability density function $f(\varepsilon)$ defined on $[0, \infty)$ and cumulative density function $F(\varepsilon)$. Denote the actual value of ε by $\tilde{\varepsilon}$. The hazard rate of ε is $\psi(\varepsilon) \equiv \frac{f(\varepsilon)}{1-\int_0^{\varepsilon} f(\eta)d\eta}$. Let $x \in X \subseteq \mathbb{R}^n$ denote the vector of state variables and define a threshold function $\Phi(x, \varepsilon) = 0$. The catastrophe occurs when $\Phi(x, \tilde{\varepsilon}) = 0$. I assume $\frac{\partial \Phi}{\partial x_i} \geq 0,\ i = 1, ..., n$ and $\frac{\partial \Phi}{\partial \varepsilon} \leq 0$: higher values of the state variables bring the system 'closer' to the threshold, and higher values of $\tilde{\varepsilon}$ imply a higher threshold. Define $\phi : X \to \mathbb{R}_+$ as $\{\varepsilon : \Phi(x, \varepsilon) = 0,\ x \in X\}$. $\phi(x)$ is the value of ε such that the threshold is reached when the state variables take on value x. Because of the assumptions on the partial derivatives of Φ, $\phi'(x) \geq 0$.

Definition 5.1. *Let $x : \mathbb{R}_+ \to X$ be continuous and differentiable almost everywhere. $x(t)$ is monotonically increasing with respect to $\Phi(x(t), \varepsilon) = 0$ and ε if and only if for any t_0 and t_1 such that $t_0 < t_1$ it holds that*

$$\Phi(x(t_0), \varepsilon_0) = \Phi(x(t_1), \varepsilon_1) \Leftrightarrow \varepsilon_0 \leq \varepsilon_1$$

For trajectories of the state variables $x(t)$ that are monotonically increasing with respect to $\Phi(x(t), \varepsilon) = 0$, $\phi(x(t))$ increases over time. From here on, I restrict attention to such trajectories, as trajectories with decreasing state variables will not be optimal.

Then the occurrence time of the catastrophe τ is a Poisson process:

$$\tau \sim f\left(\phi\left(x\left(\tau\right)\right)\right)\phi'\left(x\left(\tau\right)\right)x'\left(\tau\right)$$

Nævdal (2006) models the catastrophe as a discrete jump in the state variables. He argues that this approach is more general than a discrete jump in instantaneous utility, the approach I take in this paper. The latter can always be modeled as the former, but not the other way around. When the catastrophe occurs at time τ, the jump in the state variables is given by

$$x\left(\tau^{+}\right)=Q\left(x\left(\tau^{-}\right)\right)=x\left(\tau^{-}\right)+q\left(x\left(\tau^{+}\right)\right) \tag{5.40}$$

where $x\left(\tau^{-}\right)=\lim_{t\uparrow\tau}x\left(t\right)$ and $x\left(\tau^{+}\right)=\lim_{t\downarrow\tau}x\left(t\right)$. Nævdal (2006) shows that expected discounted utility is maximized by solving the following problem

$$\begin{aligned}
&\tilde{V}\left(t,x\left(t\right)\right)=\max_{z}\mathbb{E}\left(\int_{0}^{\infty}f\left(x,z\right)e^{-\delta t}dt\right)\ s.t.\ \dot{x}=g\left(x,z\right),\ x\left(0\right)=x_{0}\\
&x\left(\tau^{+}\right)=x\left(\tau^{-}\right)+q\left(x\left(\tau^{-}\right)\right)\\
&\tau\sim\psi\left(x\left(\tau\right),z\left(\tau\right)\right)g\left(x\left(\tau\right),z\left(\tau\right)\right)\exp\left(-\int_{0}^{\tau}\psi\left(x\left(s\right)\right)g\left(x\left(s\right),z\left(s\right)\right)ds\right)
\end{aligned} \tag{5.41}$$

where we write $g\left(x,z\right)$ for $x'\left(t\right)$. The risk-augmented Hamiltonian for this problem is

$$\begin{aligned}
H\left(x,\mu,z\right)=u\left(x,z\right)+\mu g\left(x,z\right)+\psi\left(\phi\left(x\right)\right)\phi'\left(x\right)g\left(x,z\right)\\
\times\left[\tilde{V}\left(t,x+q\left(x\right)|\tau=t\right)-\tilde{V}\left(t,x\right)\right]
\end{aligned} \tag{5.42}$$

where

$$\tilde{V}\left(t,x|\tau=t\right)=\max_{z}\int_{t}^{\infty}u\left(y,z\right)e^{-\delta\left(s-t\right)}ds\ s.t.\ \dot{y}=g\left(y,z\right),\ y\left(t\right)=x \tag{5.43}$$

is the value of continuing optimally when the catastrophe occurs at time t and results in state x. For brevity, I write $(.|\tau)$ as shorthand for $(.|\tau=t)$. The post-catastrophe problem is a standard deterministic control problem with costate variables $\mu\left(s,t|\tau\right)$. Note that $\frac{\partial}{\partial x}\tilde{V}\left(t,x|\tau\right)=\mu\left(t,t|\tau\right)$ and $\frac{\partial}{\partial x}\tilde{V}\left(t,x+q\left(x\right)|\tau\right)=\left(I^{n}+q'\left(x\right)\right)\mu\left(t,t|\tau\right)$, where I^{n} is the n-dimensional identity matrix and $q'\left(x\right)$ is the Jacobian of $q\left(x\right)$. Lastly, $J\left(t,x\right)$

in (5.42) is

$$\tilde{V}(t,x) = \max_{z} \mathbb{E}\left(\int_t^{\infty} u(y,z)\, e^{-\delta(s-t)} ds\right) \;\; s.t.\; \dot{x} = g(y,z)\,,\;\; y(0) = x$$
$$x\left(\tau^{+}\right) = x\left(\tau^{-}\right) + q\left(x\left(\tau^{-}\right)\right)$$
$$\tau \sim \psi\left(x(\tau), z(\tau)\right) g\left(x(\tau), z(\tau)\right) \exp\left(-\int_0^{\tau} \psi\left(x(s)\right) g\left(x(s), z(s)\right) ds\right) \tag{5.44}$$

The differential equation for $\tilde{v} = \tilde{V}(t, x(t))$ is then (see the Appendix in Nævdal (2006))

$$\dot{\tilde{v}} = \delta\tilde{v} - u(x,z) + \psi\left(\phi(x)\right)\phi'(x)\, g(x,z)\left(\tilde{v} - \tilde{V}\left(t, x + q(x)\,|\tau\right)\right) \tag{5.45}$$

The Hamiltonian (5.42) gives rise to the following conditions

$$u = \operatorname*{argmax}_{\upsilon} H(x, \mu, \upsilon) \tag{5.46}$$
$$\dot{\mu} = \delta\mu - \frac{\partial}{\partial x} f(x,z) - \mu\frac{\partial}{\partial x} g(x,z) - \lambda(x)\left(\mu\left(t|t, x+q(x)\right)\left(I^n + q'(x)\right) - \mu\right)$$
$$- \lambda'(x)\left(\tilde{V}\left(t, x+q(x)\,|\tau\right) - \tilde{v}\right) \tag{5.47}$$

where $\lambda(x) = \psi(\phi(x))\,\phi'(x)\, g(x,z)$. Lastly, define the transversality conditions. If x is the optimal path, then for all admissible y and $\dot{y} = g(y,u)$, we must have

$$\lim_{t\to\infty} \mu e^{-\delta t}\left(y(t) - x(t)\right) \geq 0 \qquad\qquad \lim_{t\to\infty} z(t)\, e^{-\delta t} = 0 \tag{5.48}$$

Problem (5.22) has a single state variable: $x = D$. The growth rate $g(z,D)$ of the pollution stock is $z - \alpha D$ and the catastrophe hazard is $\psi(\phi(D)) = \psi(D)$. The stock does not affect utility directly, so $\mu(t|t, x+q(x)) = 0$. Because the optimal z post-catastrophe is arbitrarily large and the first generation conditions its strategy on catastrophe occurrence, the jump in the state variable $q(D)$ at time τ is $\bar{u} - \underline{u} + \delta\xi e^{\delta t'}$, where $\bar{u} = \lim_{z\to\infty} u(z)$. This ensures that post-catastrophe generations receive utility $\underline{u}$ and $\tilde{v} - \tilde{V}(t, x+q(x)\,|\tau = t) = \tilde{v} - \frac{\underline{u}}{\delta} + \xi e^{\delta t'}$. Equations (5.23) follow by substituting in (5.45) and (5.46).

Chapter 6

The management of nonrenewable resources with amenity value under hyperbolic discounting

6.1. Introduction

Some important natural resources provide a perpetual stream of benefits when left intact, but do not regenerate when economic activity directly or indirectly causes them to deteriorate. Biodiversity provides a wide range of ecosystem services and plays an important role in pharmaceutical research (Nunes and van den Bergh, 2001), and a low carbon concentration in the atmosphere ensures a hospitable climate. Both of these resources can be expended for immediate economic gain - e.g. by cutting down the habitat of an endangered species - but even small disruptions cause irreversible damage. Extinct species are gone forever, and the mean lifetime of carbon emissions from fossil fuels is in the tens of thousands of years (Archer, 2005).

The decision whether to 'consume' these resources balances immediate gains against costs that are spread over many generations. There are positive and normative reasons to use a time-declining pure rate of time preference (DPRTP) for such decisions. DPRTPs arise naturally when there is uncertainty about the exponential discount rate or when aggregating the time preferences of heterogeneous individuals (Weitzman, 2001). DPRTPs can simultaneously explain the short-term impatience implied by capital market interest rates[75] and concerns for the long-run future from stated preference surveys[76] and intro-

[75]Nordhaus (1994) argues that observed short- and medium-term interest rates reveal a high time preference, but extrapolating these rates to long horizons is not innocuous.

[76]Layton and Brown (2000) find no significant difference in willingness to pay for preventing environmental damages that occur in 60 or 150 years.

spection (Gerlagh and Liski, 2012). DPRTPs also address the normative concern that individuals are entitled to be impatient within their own lifetime, but not to similarly discount the welfare of future generations.

This paper analyzes the consumption of a nonrenewable resource with amenity value under hyperbolic discounting, a special case of DPRTP. Marginal consumption utility is decreasing and bounded; the amenity value is linearly proportional to the remaining stock;[77] the number of generations is infinite. It is well known that hyperbolic preferences are dynamically inconsistent: generation t values utility at $t+2$ relatively more vis a vis utility at $t+1$ than generation $t+1$ does (Laibson, 1998). As a result, today's generation may appreciate the resource for its ability to provide amenity value into the far future, but its valuation of the resource goes down significantly if it believes the next generation will deplete the resource, because each generation applies a high short-term discount rate. Therefore, today's generation is more inclined to conserve if it believes that future generations will also conserve, but more likely to consume if future generations will also consume.

For an intermediate range of amenity values, the intergenerational game in which regulators are sophisticated and foresee the different preferences of their successors has multiple equilibria: a Pareto-dominant one in which the resource is fully conserved, and a large number of suboptimal equilibria with positive consumption. By contrast, naive regulators who believe they can fully commit future actions always conserve: since they have no intention to consume in the future, the motive to preempt their successors' consumption does not exist for these naifs. Fully rational policies can thus lead to worse outcomes than policies that are based on naive beliefs.[78]

Still, large initial resource stocks can only be partially rather than fully exhausted in subgame perfect equilibrium. Each generation only consumes a unit of the resource if it

[77] Many of our results go through with concave amenity values, but for clarity and tractability we prefer the linear specification.

[78] A similar "sophistication effect" occurs in O'Donoghue and Rabin (1999) and Grenadier and Wang (2007). In the former, an agent decides when to perform an activity that confers an immediate reward but can only be performed once. In the latter, an investor decides when to exercise an American call option that is in the money. Sophisticated agents anticipate that their future selves will perform the activity or exercise the option too early from the point of view of the current self, and therefore perform the activity suboptimally early in equilibrium compared to their naive counterparts. In contrast to the current paper, there is no multiplicity of equilibria. The equilibrium in O'Donoghue and Rabin (1999) is unique because the number of periods is finite; in Grenadier and Wang (2007) because the volatility in the value of the underlying asset ensures the option is always exercised in finite time. For an application of the model in Grenadier and Wang (2007) to forest economics, see Di Corato (2012).

believes that subsequent generations will otherwise consume this unit sufficiently quickly. But future generations' marginal propensity to consume cannot be high for all levels of the resource stock - otherwise, consumption in some future periods would go to infinity when the initial resource stock is very large, and the marginal utility from consumption in those periods would fall short of the welfare gain from enjoying the amenity.

Our results suggest that the management of nonrenewable resources with amenity values has flavours of a stag hunt game. Each generation may prefer to see the resource conserved indefinitely, but only implements a conservationist policy if it believes that the resource will not be used up in the future. There may thus be a role for policies that help generations to coordinate on the Pareto-dominant conservation equilibrium. In line with experimental evidence that framing is important for cooperation in stag hunt games (Ellingsen et al., 2012), designating the natural resource as a nature reserve can help future generations to coordinate on the superior conservation equilibrium. Such a policy may also have a commitment role by creating legal hurdles or legislative costs for future exploitation.

Interestingly, our result that resource depletion is more rapacious under sophisticated than under naive beliefs contrasts with findings from previous work on renewable resources (Karp, 2005; Fuji and Karp, 2006; Hepburn et al., 2010).[79] In case of an nonrenewable resource with amenity value, naive regulators understand that current consumption irreversibly reduces long-run utility, tempering their impulse to consume. Resource regeneration severs this link between current-period and long-run utility however. A naive regulator that manages a renewable resource with amenity value may ultimately aim for the steady state corresponding to his long-run discount rate, but does not realize that his successors will not replenish the stock to this level. Sophisticated regulators on the other hand reduce their consumption to compensate for the overconsumption of their successors. This intuition suggests that increased resilience need not be a blessing when an ecosystem is managed by naive policy makers, because resource regeneration instills in naifs an unwarranted optimism regarding future stocks. In the Appendix, we

[79]Karp (2005) considers a model in which consumption contributes to a pollution stock that reduces utility. In each period, a fraction of the stock decays naturally. Karp shows that regulators with full commitment power start off with higher consumption than in Markov equilibrium, but stabilize the stock at a lower level. It follows that a sequence of naive regulators would reach a higher steady-state stock than in Markov equilibrium. Hepburn et al. (2010) show that a sequence of naive regulators may eventually extinguish a fishery stock (without amenity value), even though early regulators do not intend such an eventual collapse.

provide a numerical example in which steady state welfare and welfare under stationary discounting are lower under naive beliefs when the resource regenerates, compared to the case of a nonrenewable resource.

6.2. Model

Consider a sequence of generations living in periods $t = 1, 2, ..., \infty$, with each generation represented by a single agent. The economy contains a nonrenewable natural resource that has both consumption and amenity value. Generation t's marginal consumption utility is bounded from above by ϕ, and instantaneous felicity is linear in the remaining stock with slope θ:

$$\begin{aligned} v_t =& u(q_t) + \theta (S_t - q_t) \\ & u'(.) > 0,\ u''(.) < 0 \\ & 0 = \lim_{q \to \infty} u'(q) < \theta < \phi \equiv u'(0) < \infty \end{aligned} \tag{6.1}$$

subject to stock dynamics

$$S_{t+1} = S_t - q_t \tag{6.2}$$

The representative agent in period t has hyperbolic preferences over its own felicity and the felicity of all subsequent generations. The welfare of generation t, given a sequence $\{q_{t+\sigma}, S_{t+\sigma}\}_{\sigma=0}^{\infty}$, is

$$w_t\left(\{q_{t+\sigma}, S_{t+\sigma}\}_{\sigma=0}^{\infty}\right) = v_t + \beta \sum_{\tau=1}^{\infty} \delta^{\tau} v_{t+\tau},\ \ \beta, \delta \leq 1 \tag{6.3}$$

When $\beta < 1$, decision-makers are more impatient with respect to short-term outcomes than with respect to long-term outcomes. Preferences are time-inconsistent in this case: the relative importance of v_t compared to $v_{t+\tau}$ is higher from the perspective of agent t than from agent $t-1$ for all τ.

When the resource's amenity value is sufficiently high, full conservation ($q_t = 0\ \forall\ t$) Pareto-dominates all other consumption paths. That is, each generation prefers full conservation to any other consumption path.

Proposition 6.1. *If* $\phi < \theta \left[\frac{\beta}{1-\delta} + 1 - \beta\right]$*, then*

$$w_t\left(\{q_{t+\sigma} = 0\}_{\sigma=0}^{\infty}\right) > w_t\left(\left\{q'_{t+\sigma}\right\}_{\sigma=0}^{\infty}\right) \;\; \forall\; q' \neq q$$

As a result, naive policy makers fully conserve the resource under the parameter values in Proposition 6.1. Sophisticated regulators play a game against their successors. In Markov perfect equilibrium (MPE), the current generation's strategy $q_t = \zeta(S_t)$ depends only on the current stock and is a best response if future generations will play $\zeta(S_{t+\tau}) \;\forall\; \tau$:

$$\begin{aligned} \zeta(S_t) = \underset{q}{\operatorname{argmax}} \Bigg\{ & u(q) + \theta(S_t - q) \\ & + \beta \sum_{\tau=1}^{\infty} \delta^{\tau}\left[u(\zeta(S_{t+\tau})) + \theta(S_{t+\tau} - \zeta(S_{t+\tau}))\right] \Bigg\} \\ & \text{where } S_{t+\tau} \text{ is given recursively by } S_t - \sum_{\sigma=0}^{\tau-1} \zeta(S_{t+\sigma}) \end{aligned} \tag{6.4}$$

Initially, we only exclude mixed-strategy equilibria; later, we confine ourselves to everywhere differentiable strategies.

Proposition 6.2. *If* $\phi \leq \theta \left[\frac{\beta}{1-\delta} + 1 - \beta\right]$*, then* $\zeta(S) = 0 \;\forall\; S$ *is a Markov perfect equilibrium (MPE). If* $\phi > \theta \left[\frac{\beta}{1-\delta} + 1 - \beta\right]$*, then* $\zeta(S') = 0 \;\forall\; S' \leq S$ *cannot be a MPE.*

Proof. When all future generations choose $\zeta(S) = 0$, the result follows directly from the first line of (6.7). Q.E.D.

There also exist equilibria with positive consumption. When preferences are inconsistent ($\beta < 1$), the current generation discounts consumption of the resource in the immediate future at a comparatively high rate vis a vis the distant future. If the current generation believes that the next generation(s) will consume the resource, the long-term amenity value of the resource will not be realized. The current generation can then only influence whether the resource will be consumed today or in the near future. Because the near-term discount rate is relatively high, the current generation may prefer to consume the resource itself rather than having it consumed by its immediate descendants.

Resource depletion can thus become a self-fulfilling belief, even when full conservation is the Pareto-dominant equilibrium - as we illustrate in the next Example. Denote $\bar{S}$ as the lowest S for which $\zeta(S) = S$ is an equilibrium, so that $\theta + \beta\delta\phi = u'\left(\bar{S}\right)$.

Example 6.1. *Let $S \leq \bar{S}$ with $\bar{S}$ and*

$$\frac{\theta}{1-\beta\delta} < \phi < \theta\left(\frac{\beta}{1-\delta} + 1 - \beta\right)$$

Then $\zeta(S) = S$ and $\zeta(S) = 0$ are both MPEs.

Dynamic games with hyperbolic discounting have a wide range of equilibria, especially if we allow discontinuous and nondifferentiable strategies (Krusell et al., 2002; Krusell and Smith, 2003). For example, we can obtain a new MPE by modifying an existing MPE such that $\zeta(S) = 0$ for S larger than some S'. Conservation is an equilibrium for large S per Proposition 6.2, and the original strategy remains an equilibrium for small S.

Corollary 6.1. *If $\zeta(S)$ is a MPE, then*

$$\tilde{\zeta}(S) = \begin{cases} 0 & \text{for } S \geq S' \\ \zeta(S) & \text{for } S < S' \end{cases}$$

is also a MPE $\forall\ S'$.

Each equilibrium strategy $\zeta(S)$ can also be used to formulate a new equilibrium for a different initial stock. To see, this, suppose we increase the initial stock by Δ units. If each generation behaves as if these Δ will be preserved forever, no generation has an incentive to modify its behaviour and consume these units. The extra units then act as a lump-sum benefit of $\theta\left(\frac{\beta}{1-\delta} + 1 - \beta\right)$ to each generation, and $\tilde{\zeta}(S+\Delta) = \zeta(S)$ is an equilibrium.[80]

Proposition 6.3. *Let $\zeta(S)$ be a MPE and define $\tilde{\zeta}(S) : \tilde{\zeta}(S+\Delta) = \zeta(S)\ \forall\ S$. Then $\tilde{\zeta}(S)$ is also a MPE.*

[80] If the amenity value is subject to decreasing returns, a modified version of Proposition 6.3 holds for small stocks.

Proposition 6.3 implies that equilibria can also be defined in terms of cumulative resource use, rather than in terms of the initial stock. By Example 6.1, there exist equilibria in which a small initial resource stock, of say 1 unit, will be completely exhausted. Then there also exists an equilibrium in which an initial stock of 101 units is consumed until there are 100 units remaining, which are then conserved indefinitely.

Now confine attention to policy rules that are everywhere differentiable. We also impose $\zeta'(S) \leq 1$, which seems a 'reasonable' property for an equilibrium. Hereafter, we call this an 'admissible' equilibrium. As it turns out, we can straightforwardly rule out $\zeta'(S) = 1$.

Lemma 6.1. $\zeta'(S) \neq 1 \;\forall\; S \geq \bar{S}$

Generations with positive consumption balance two considerations: the marginal increase in instantaneous felicity must equal the marginal reduction in discounted future felicities and amenity values. If the reaction function were to have slope 1, the latter term would be equal for two sufficiently close values of the stock, but the former would not. Therefore, $\zeta'(S) = 1$ implies a violation of the first-order condition.

We already established that whenever conservation is Pareto-dominant, conservation is an equilibrium. When the consumption value of the resource is sufficiently low, full conservation is the only MPE:

Proposition 6.4. *For $\phi < \frac{\theta}{1-\beta\delta}$, as well as for $\phi = \frac{\theta}{1-\beta\delta}$ and $S > \bar{S}$, $\zeta(S) = 0$ in all admissible equilibria.*

At the threshold value of ϕ, generation t is indifferent between enjoying the amenity value of a marginal resource unit for one period and having the unit consumed by generation $t+1$ and consuming the unit in period t. This threshold for ϕ coincides with the value that permits the exhaustion of small resource stocks from Example 6.1 and, by extension, some discontinuous consumption equilibria from Proposition 6.3.

Proposition 6.5 shows that consumption must be equal to zero for some values of the stock in any equilibrium.

Proposition 6.5. *(i) If $\phi < \theta\left(\frac{\beta}{1-\delta} + 1 - \beta\right)$, then there exists an S^* such that $\zeta(S^*) = 0$.*

(ii) For ϕ in a right neighbourhood of $\frac{\theta}{1-\beta\delta}$, there exists an $\tilde{S}$ such that all admissible equilibria satisfy $\zeta(S) = 0 \;\forall\; S > \tilde{S}$.

When ϕ is in the range for which conservation is Pareto-dominant but positive consumption is an equilibrium strategy, generation t only chooses to consume a marginal resource unit in equilibrium if it believes that subsequent generations will otherwise consume this unit sufficiently quickly. More specifically, because $\delta < 1$ and $\zeta'(S) < 1$, generation t only consumes a marginal unit if future generations will otherwise consume this unit sufficiently quickly. Future generations' marginal propensity to consume cannot be large for all levels of the resource stock however - otherwise, resource consumption tends to infinity in some future periods and marginal utility tends to zero when S becomes very large. This would contradict the requirement that marginal consumption utility always exceeds the current-period amenity value θ. Therefore, there is some value of the resource stock S^* for which future generations will conserve a marginal unit for a large number of periods, and for which the current generation hence chooses $\zeta(S^*) = 0$. At S^*, the diminishing marginal utility of future generations acts as a coordination device.

By a similar argument, large initial stocks are never fully exhausted in equilibrium.

Proposition 6.6. *Let the initial stock S_0 be sufficiently large. Then in any admissible equilibrium,* $\sum_{t=0}^{\infty} \zeta(S_t) < S_0$.

Propositions 6.3 and 6.6 indicate that the steady-state value of the resource stock may depend on the initial stock. Large initial stocks are not reduced below some S^* in continuously differentiable equilibria, but if the initial stock equals S^*, there exist equilibria with $\zeta(S^*) > 0$.

6.3. A regeneration paradox: a more resilient resource reduces welfare

The previous section demonstrates that a nonrenewable resource with amenity value is depleted less rapaciously by naive than by sophisticated regulators: a sophisticated regulator may consume the resource today in order to preempt his successors, whereas a naif does not plan to consume in the future and is hence more conservationist. The suboptimal outcomes (e.g. undersaving or overconsumption) that often arise in hyperbolic discounting models with naive regulators do not occur: because the resource does not regenerate, the naif understands that he cannot compensate for splurging today with future restraint. When the resource does regenerate, a naive regulator may be tempted

Table 6.1: Parameter values

	Renewable	Nonrenewable
a	2.3	2.3
b	1	1
θ	0.185	0.185
γ	0.06	0
S_0	0.9	0.9
β	0.72	0.72
δ	0.95	0.95

to choose a high level of consumption today, mistakenly believing that the resource will regrow to his desired long-run level in the future.

In this section, we explore the conjecture that resource regeneration decreases welfare under naive beliefs, because it exacarbates the naif's misperception regarding future consumption and future resource stocks. We provide a numerical example in which steady state welfare and welfare under constant discounting are lower under naive beliefs when the resource is renewable, compared to when it is nonrenewable. Let the growth of the resource be given by a Rickert curve

$$S_{t+1} = S_t e^{\gamma(1-S_t)} - q_t \tag{6.5}$$

where a nonrenewable resource is described through $\gamma = 0$, and consider the quadratic utility function

$$v_t = aq_t - \frac{1}{2}bq_t^2 + \theta S_{t+1} \tag{6.6}$$

Table 6.1 lists the parameter values. They are such that $\frac{\theta}{1-\beta\delta} < \phi < \theta\left(\frac{\beta}{1-\delta} + 1 - \beta\right)$, meaning that a naive regulator would conserve indefinitely if the resource is nonrenewable. When consumption is positive, we must have

$$v_q\left(q_t, S_t\right) - \beta\delta V'\left(S_{t+1}\right) = 0 \Leftrightarrow a - bq_t - \theta - \beta\delta V'\left(S_{t+1}\right) = 0$$

where V is the value function of the time-consistent problem with constant discount factor δ. We estimated the value function with the `CompEcon` package from Miranda and Fackler (2002), and its numerical derivative using a two-sided difference approach with 3000 grid points. Figure 6.1 depicts the value function for different levels of the stock. When the resource is nonrenewable, the value function is linear in the stock. In case of a renewable resource, the marginal value of the resource is low for large and intermediate stocks. Even for some S below the value that maximizes the sustainable yield ($S = 0.4962$), $V'(S)$ is smaller when $\gamma = 0.06$ than when $\gamma = 0$: because the resource can be replenished in the future, the loss from present consumption is not as large as in the case of a nonrenewable resource.

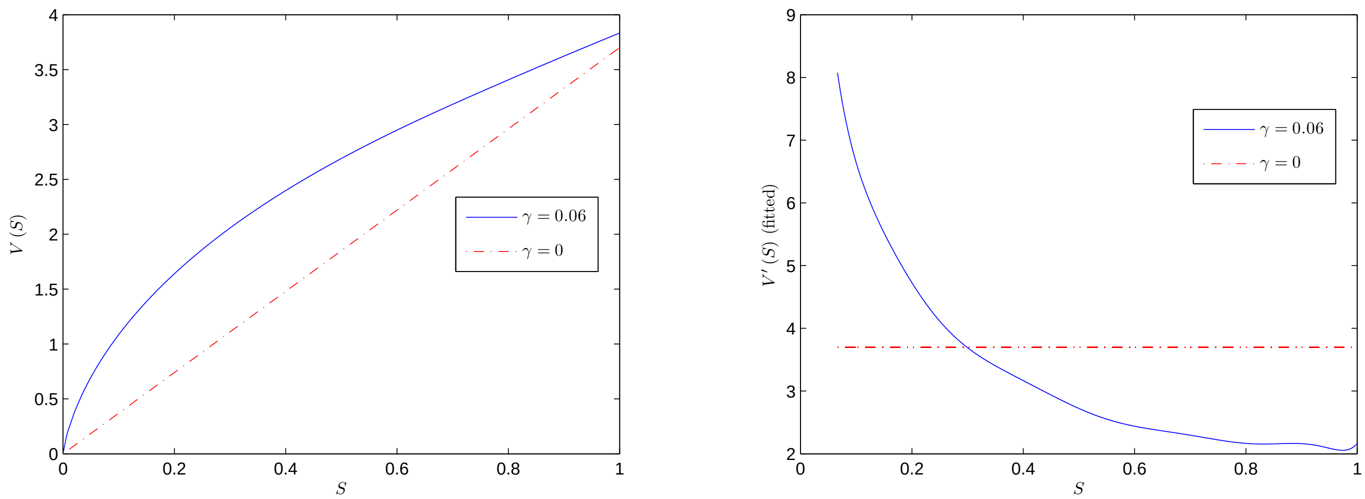

Figure 6.1: Value function (left) and marginal value (right) for renewable and nonrenewable resource

Table 6.2 shows the evolution of consumption q_t, the resource stock S_t, felicity v_t and welfare w_t. A nonrenewable resource is always conserved, so the values in the right half of the table are constant over time. For high levels of the stock, the marginal value of the resource is relatively low when the resource is renewable, so initial consumption is very high in this case. Long-run consumption is positive when $\gamma = 0.06$, but the steady-state stock is less than half of the initial value. As a result, the first generation's welfare is slightly higher when the resource is renewable than when it is nonrenewable - owing to the high first-period consumption - but all subsequent generations are worse off compared to the nonrenewable-resource benchmark. The often-applied welfare criterion of the felicity stream with stationary discounting, $\sum_{t=1}^{\infty} \delta^{t-1} v_t$, is also lower when the

Table 6.2: Naive consumption, felicity and welfare paths for renewable and nonrenewable resource

	Renewable				Nonrenewable			
t	q_t	S_t	v_t	w_t	q_t	S_t	v_t	w_t
1	0.360	0.9	0.864	2.590	0	0.9	0.167	2.444
2	0.109	0.545	0.329	1.909	0	0.9	0.167	2.444
3	0.038	0.451	0.166	1.710	0	0.9	0.167	2.444
4	0.021	0.428	0.125	1.660	0	0.9	0.167	2.444
5	0.016	0.423	0.115	1.648	0	0.9	0.167	2.444
6	0.015	0.421	0.113	1.645	0	0.9	0.167	2.444
7	0.015	0.421	0.112	1.645	0	0.9	0.167	2.444
8	0.015	0.421	0.112	1.644	0	0.9	0.167	2.444
9	0.015	0.421	0.112	1.644	0	0.9	0.167	2.444
10	0.015	0.421	0.112	1.644	0	0.9	0.167	2.444

This table contains q_t, S_t, v_t and w_t when the regulator in control at time t acts as if he can commit all future consumption. The welfare function is given by (6.3) with functional forms (6.5) and (6.6). Parameter values are in Table 6.1. A time-consistent regulator with discount factor δ would stabilize the stock at $S = 0.766$.

resource is renewable (3.262) than when it is nonrenewable (3.340).

6.4. Conclusion

Amenity values provide a perpetual flow of services that are valued by policy makers that have a special concern for immediate consumption, but discount the distant future at a low rate. Conservation efforts crucially depend on the current policy maker's confidence that natural resource will remain intact in the future, and not be depleted for immediate economic gains in the medium term. Sophisticated regulators understand that they cannot control future policies, and may thus be inclined to consume the resource before their successors will. Public commitment or coordination devices, such as the establishment of nature reserves, may remedy the coordination problem between generations.

Naive regulators do not contend with future depletion, and are thus more likely to follow a conservationist policy. The naif's shortsightedness does not lead to suboptimal policies precisely because the resource is nonrenewable: the naive regulator values the far future, and understands that future well-being is irreversibly compromised by current depletion. This paper has raised the possibility that resource regeneration may be a curse for these naifs, as it could lead to a mistaken belief that future replenishment can compensate for today's consumption.

6.A. Appendix: Proofs

Proof of Proposition 6.1

Proof. It is sufficient to show that a generation's welfare decreases when it increases consumption itself and when its descendants increase consumption.

$$\begin{aligned}
\frac{\partial w_t}{\partial q_t} &= -\theta\left(1+\beta\sum_{\tau=1}^{\infty}\delta^{\tau}\right)+\phi = -\theta\left(\frac{\beta}{1-\delta}+1-\beta\right)+\phi < 0 \\
\left.\frac{\partial w_t}{\partial q_{t+\sigma}}\right|_{\sigma>0} &= \beta\delta^{\sigma}\left(-\theta\sum_{\sigma'=\sigma+1}^{\infty}\delta^{\sigma'-\sigma-1}+\phi\right) = \beta\delta^{\sigma}\left[-\frac{\theta}{1-\delta}+\phi\right] \\
&< \beta\delta^{\sigma}\left(-\theta\left(\frac{\beta}{1-\delta}+1-\beta\right)+\phi\right) < 0 \qquad (6.7)
\end{aligned}$$

Q.E.D.

Proof of Example 6.1

Proof. By Proposition 6.2, the second inequality guarantees that $\zeta(S) = 0$ is an MPE. Now suppose that the current generation believes that its successors will choose $\zeta(S) = S$. Then it will also want to choose $\zeta(S) = S$ if the marginal welfare of consuming the last unit of resource itself exceeds the marginal welfare of having the last unit consumed by the next generation, that is if

$$\theta + \beta\delta\phi < u'(S) \le \phi$$

A solution only exists if the current generation's welfare gain from consuming the resource today instead of tomorrow exceeds the per-period amenity value ($\phi(1-\beta\delta) > \theta$).

Q.E.D.

Proof of Proposition 6.3

Proof. From the definition of an equilibrium

$$\begin{aligned}\zeta(S_t) = &\operatorname*{argmax}_q \{u(q) + \theta(S_t - q)\\ &+\beta \sum_{\tau=1}^{\infty} \delta^\tau [u(\zeta(S_{t+\tau})) + \theta(S_{t+\tau} - \zeta(S_{t+\tau}))]\Bigg\}\\ = &\operatorname*{argmax}_q \{u(q) + \theta(S_t - q)\\ &+\beta \sum_{\tau=1}^{\infty} \delta^\tau [u(\zeta(S_{t+\tau})) + \theta(S_{t+\tau} - \zeta(S_{t+\tau}))]\\ &+\theta\Delta\left[\frac{\beta}{1-\delta} + 1 - \beta\right]\Bigg\}\\ &\text{where } S_{t+\tau} \text{ is given recursively by } S_t - \sum_{\sigma=0}^{\tau-1} \zeta(S_{t+\sigma})\end{aligned}$$

Suppose that all generations use $\tilde{\zeta}(S_t + \Delta) = \zeta(S_t)\ \forall\ S_t$. Then for all S_t,

$$\begin{aligned}&u(q) + \theta(S_t - q) + \beta \sum_{\tau=1}^{\infty} \delta^\tau [u(\zeta(S_{t+\tau})) + \theta(S_{t+\tau} - \zeta(S_{t+\tau}))]\\ &+\theta\Delta\left[\frac{\beta}{1-\delta} + 1 - \beta\right] =\\ &u(q) + \theta(S_t + \Delta - q) + \beta \sum_{\tau=1}^{\infty} \delta^\tau \left[u\left(\zeta\left(\tilde{S}_{t+\tau}\right)\right) + \theta\left(\tilde{S}_{t+\tau} - \zeta\left(\tilde{S}_{t+\tau}\right)\right)\right]\\ &\text{where } \tilde{S}_{t+\tau} \text{ is given recursively by } S_t + \Delta - \sum_{\sigma=0}^{\tau-1} \zeta\left(\tilde{S}_{t+\sigma}\right)\end{aligned} \tag{6.8}$$

Because $q = \zeta(S_t)$ maximizes the left hand side of (6.8), $q = \tilde{\zeta}(S_t + \Delta) = \zeta(S_t)$ maximizes the right hand side. Then $\tilde{\zeta}(S_t + \Delta)$ satisfies the conditions of an equilibrium.

Q.E.D.

Proof of Lemma 6.1

Proof. From (6.4), generation t's first order condition when $\zeta(S_t) > 0$ is

$$\frac{\partial w_t}{\partial q_t} = u'(q_t) - \theta - \beta \sum_{\tau=1}^{\infty} \left\{ \delta^{\tau} \underbrace{\prod_{\sigma=1}^{\tau-1} [1 - \zeta'(S_{t+\sigma})]}_{I} [u'(q_{t+\tau}) \zeta'(S_{t+\tau}) + \theta (1 - \zeta'(S_{t+\tau}))] \right\} = 0 \tag{6.9}$$

If generation t conserves an additional resource unit, term I indicates the fraction of this unit that survives to time $t+\tau$. The subsequent term in square brackets denotes how the addition of one resource unit at time $t+\tau$ affects the felicity of generation $t+\tau$, which from generation t's perspective is discounted at rate $\beta\delta^{\tau-t}$. When there is an interior solution such that (6.9) holds, $\zeta'(S^*) = 1$ implies that (6.9) must hold for S^* and $S^*+\epsilon$ (ϵ small). But the third term in (6.9) is equal for the $S_{t+1} = S^* - \zeta(S^*) = S^* + \epsilon - \zeta(S^* + \epsilon)$, whereas marginal utility is different for $\zeta(S^*)$ and $\zeta(S^* + \epsilon) = \zeta(S^*) + \epsilon$. Therefore, equation (6.9) cannot hold for both S^* and $S^* + \epsilon$, so we cannot have $\zeta'(S^*) = 1$ for any S^*. Q.E.D.

Proof of Proposition 6.4

Proof. The Euler equation reads

$$
\begin{aligned}
u'(q_t) - \theta &= \beta \sum_{\tau=1}^{\infty} \left\{ \delta^{\tau} \prod_{\sigma=1}^{\tau-1} \left[1 - \zeta'(S_{t+\sigma})\right] \left[u'(\zeta(S_{t+\tau}))\zeta'(S_{t+\tau}) + \theta(1 - \zeta'(S_{t+\tau}))\right] \right\} \\
&= \beta\delta \left(\theta + \sum_{\tau=2}^{\infty} \left\{ \delta^{\tau-1} \prod_{\sigma=1}^{\tau-1} \left[1 - \zeta'(S_{t+\sigma})\right] \left[u'(\zeta(S_{t+\tau}))\zeta'(S_{t+\tau}) \right.\right.\right. \\
&\quad \left.\left. + \theta(1 - \zeta'(S_{t+\tau}))\right]\right\} + \zeta'(S_{t+1}) \left(u'(\zeta(S_{t+1})) - \theta - \right. \\
&\quad \left.\left. \sum_{\tau=2}^{\infty} \left\{ \delta^{\tau-1} \prod_{\sigma=1}^{\tau-1} \left[1 - \zeta'(S_{t+\sigma})\right] \left[u'(\zeta(S_{t+\tau}))\zeta'(S_{t+\tau}) + \theta(1 - \zeta'(S_{t+\tau}))\right] \right\} \right) \right) \\
&= \beta\delta \left(\theta + \sum_{\tau=2}^{\infty} \left\{ \delta^{\tau-1} \prod_{\sigma=1}^{\tau-1} \left[1 - \zeta'(S_{t+\sigma})\right] \left[u'(\zeta(S_{t+\tau}))\zeta'(S_{t+\tau}) \right.\right.\right. \\
&\quad \left.\left. + \theta(1 - \zeta'(S_{t+\tau}))\right]\right\} - \zeta'(S_{t+1}) \left((1 - \beta) \right. \\
&\quad \left.\left. \sum_{\tau=2}^{\infty} \left\{ \delta^{\tau-1} \prod_{\sigma=1}^{\tau-1} \left[1 - \zeta'(S_{t+\sigma})\right] \left[u'(\zeta(S_{t+\tau}))\zeta'(S_{t+\tau}) + \theta(1 - \zeta'(S_{t+\tau}))\right] \right\} \right) \right) \\
&= \beta\delta \left(\theta + \frac{1}{\beta}\left(u'(\zeta(S_{t+1})) - \theta\right) - \zeta'(S_{t+1}) \left((1 - \beta)\frac{1}{\beta}\left(u'(\zeta(S_{t+1})) - \theta\right) \right) \right) \\
&= \delta\left[\beta\theta + \left(1 - (1 - \beta)\zeta'(S_{t+1})\right)\left(u'(\zeta(S_{t+1})) - \theta\right)\right] \qquad (6.10)
\end{aligned}
$$

The third equality follows from the observation that by generation $t+1$'s FOC,

$$
\sum_{\tau=2}^{\infty} \left\{ \delta^{\tau-1} \prod_{\sigma=1}^{\tau-1} \left[1 - \zeta'(S_{t+\sigma})\right] \left[u'(\zeta(S_{t+\tau}))\zeta'(S_{t+\tau}) + \theta(1 - \zeta'(S_{t+\tau}))\right] \right\} = \frac{1}{\beta}\left(u'(\zeta(S_{t+1})) - \theta\right)
$$

By $\zeta'(S) \leq 1$, from (6.10) it follows that

$$
u'(q_t) \geq \frac{\theta}{1 - \beta\delta} \qquad (6.11)
$$

But by the assumptions in the Proposition text, $u'(q_t) \leq \phi < \frac{\theta}{1-\beta\delta}$, contradicting (6.11). Therefore, we cannot have $\zeta(S) > 0$ when $\phi < \frac{\theta}{1-\beta\delta}$. By Lemma 6.1, $\zeta'(S)$ is strictly smaller than one when $S > \bar{S}$. In this case, (6.11) holds with strict inequality, contradicting $u'(q_t) \leq \phi = \frac{\theta}{1-\beta\delta}$. Q.E.D.

Proof of Proposition 6.5

Proof. (i) Suppose that $\zeta(S) > 0 \;\forall\; S$. Equation (6.9) implies

$$\begin{aligned}&\phi - \theta - \beta \sum_{\tau=1}^{\infty} \left\{ \delta^{\tau} \prod_{\sigma=1}^{\tau-1} \left[1 - \zeta'(S_{t+\sigma})\right] \left[u'(q_{t+\tau})\,\zeta'(S_{t+\tau}) + \theta\left(1 - \zeta'(S_{t+\tau})\right)\right] \right\} > \\ &u'(\zeta(S)) - \theta - \beta \sum_{\tau=1}^{\infty} \left\{ \delta^{\tau} \prod_{\sigma=1}^{\tau-1} \left[1 - \zeta'(S_{t+\sigma})\right] \left[u'(q_{t+\tau})\,\zeta'(S_{t+\tau}) + \theta\left(1 - \zeta'(S_{t+\tau})\right)\right] \right\} = 0\end{aligned} \tag{6.12}$$

Consider some very large T. Given that generation t prefers to conserve a marginal resource unit if future generations will conserve this unit indefinitely ($\phi < \theta\left(\frac{\beta}{1-\delta} + 1 - \beta\right)$), it can only be optimal to choose $q_t > 0$ if at least a fraction ϵ_T bounded away from zero of this unit will have been consumed before time $t+T$,[81] i.e. if

$$\sum_{\tau=1}^{T} \left\{ \prod_{\sigma=1}^{\tau-1} \left[1 - \zeta'(S_{t+\sigma})\right] \zeta'(S_{t+\tau}) \right\} \geq \epsilon_T \tag{6.13}$$

By (6.13), $\frac{\partial \sum_{\tau=1}^{T} q_{t+\tau}}{\partial S} \geq \epsilon_T \;\forall\; S$. But resource use is bounded in each period: the marginal utility of resource consumption $u'(q_t)$ must always exceed the current-period amenity value θ because of the assumption that $\zeta'(S) \leq 1$ throughout. If (6.13) were to hold for all S, resource use in some periods would go to infinity when S becomes arbitrarily large. Therefore, there must exist an S^* such that $\zeta(S^*) = 0$.

(ii) For ϕ in a right neighbourhood of $\frac{\theta}{1-\beta\delta}$ and S_t sufficiently large, $\rho(S_t) > 0$ requires that $\rho'(S_{t+\tau}) \approx 1 \;\forall\; \tau \leq T$ for some large T by (6.9) and (6.10). Then analogously to (6.11), the Euler equation stipulates $u'(\zeta(S_{t+\tau})) \approx \phi \approx \frac{\theta}{1-\beta\delta}$. Totally differentiating (6.10), we have

$$\begin{aligned}u''(\zeta(S_t))\overbrace{\zeta'(S_t)}^{\approx 1} = &-\delta(1-\beta)\,\zeta''(S_{t+1})\left(u'(\zeta(S_{t+1})) - \theta\right) + \\ &\delta\left[\left(1 - (1-\beta)\,\zeta'(S_{t+1})\right) u''(\zeta(S_{t+1}))\,\zeta'(S_{t+1})\right]\end{aligned} \tag{6.14}$$

[81] We can limit attention to what happens before $t+T$ because of the assumption that $\zeta'(S) \leq 1$; otherwise, the divergence of $\zeta'(S)$ could theoretically outweigh the discount factor's convergence to zero.

Because $\zeta(S_t) \approx \zeta(S_{t+1})$ and $\zeta'(S_t) \approx \zeta'(S_{t+1}) \approx 1$, (6.14) is approximately equivalent to

$$\underbrace{u''(\zeta(S_t))}_{<0}\underbrace{\zeta'(S_t)}_{\approx 1}\underbrace{(1-\delta+(1-\beta)\zeta'(S_{t+1}))}_{>0} = -\delta(1-\beta)\underbrace{\zeta''(S_{t+1})}_{\approx 0}(u'(\zeta(S_{t+1}))-\theta) \tag{6.15}$$

which yields a contradiction. Therefore, we cannot have $\zeta(S) > 0$ for large S when ϕ is in a right neighbourhood of $\frac{\theta}{1-\beta}$. Q.E.D.

Proof of Proposition 6.6

Proof. If $\zeta(S) > 0 \ \forall \ S \leq S_0$ for all sufficiently large S_0, then analogous to the proof of Proposition 6.4, eqn (6.12) must hold for all sufficiently large S. But then $\sum_{t=1}^{T} q_t$ (T large) becomes arbitrarily large when S_0 is sufficiently large, which violates the boundedness of $\zeta(S)$. Therefore, for large S_0 there must exist an $S^* < S_0$ such that $\zeta(S^*) = 0$. It follows automatically that $\sum_{t=0}^{\infty} \zeta(S_t) < S_0$. Q.E.D.

Bibliography

Alvarez-Cuadrado, F. and N. V. Long (2009). A mixed Bentham-Rawls criterion for intergenerational equity: Theory and implications. Journal of Environmental Economics and Management 58(2), 154–168.

Archer, D. (2005). Fate of fossil fuel CO2 in geologic time. Journal of Geophysical Research 110, C09S05.

Babiker, M. H. (2005). Climate change policy, market structure, and carbon leakage. Journal of International Economics 65(2), 421–455.

Barrett, S. (1994). Self-enforcing international environmental agreements. Oxford Economic Papers 46, 878–894.

Blumsack, S. A., J. Apt, and L. B. Lave (2006). Lessons from the failure of U.S. electricity restructuring. The Electricity Journal 19(2), 15–32.

Borenstein, S. and J. Bushnell (2000). Electricity restructuring: Deregulation or reregulation? Regulation 23(2), 46–52.

Bowen, H. P., E. E. Leamer, and L. Sveikauskas (1987). Multicountry, multifactor tests of the factor abundance theory. American Economic Review 77(5), 791–809.

BP (2010). Statistical review of world energy 2010.

Buchholz, W. and M. Schymura (2012). Expected utility theory and the tyranny of catastropic risks. Ecological Economics 77, 234–239.

Burniaux, J.-M. and J. Oliveira Martins (2012). Carbon leakages: A general equilibrium view. Economic Theory 49(2), 473–495.

Chakravorty, U., B. Magné, and M. Moreaux (2006). A Hotelling model with a ceiling on the stock of pollution. Journal of Economic Dynamics and Control 30(12), 2875–2904.

Chakravorty, U., M. Moreaux, and M. Tidball (2008). Ordering the extraction of polluting nonrenewable resources. American Economic Review 98(3), 1128–1144.

Charpentier, A. D., J. A. Bergerson, and H. L. MacLean (2009). Understanding the Canadian oil sands industry's greenhouse gas emissions. Environmental Research Letters 4(1), 014005.

Chichilnisky, G. (1996). An axiomatic approach to sustainable development. Social Choice and Welfare 13(2), 231–257.

Cho, W. G., K. Nam, and J. A. Pagan (2004). Economic growth and interfactor/interfuel substitution in Korea. Energy Economics 26(1), 31–50.

Crafts, N. and A. Mulatu (2005). What explains the location of industry in Britain, 1871-1931? Journal of Economic Geography 5(4), 499–518.

Cropper, M. L. (1979). Regulating activities with catastrophic environmental effects. Journal of Environmental Economics and Management 3(1), 1–15.

Dasgupta, P. (2012). Time and the generations. In R. W. Hahn and A. Ulph (Eds.), Climate Change and Common Sense: Essays in Honour of Tom Schelling. Oxford: Oxford University Press.

Dasgupta, P., R. J. Gilbert, and J. E. Stiglitz (1983). Strategic considerations in invention and innovation: The case of natural resources. Econometrica 51(5), 1439–1448.

Davis, D. R., D. E. Weinstein, S. C. Bradford, and K. Shimpo (1997). Using international and Japanese regional data to determine whether the factor abundance theory of trade works. American Economic Review 87(3), 421–446.

Depledge, J. (2008). Striving for no: Saudi Arabia in the climate change regime. Global Environmental Politics 8(4), 9–35.

Di Corato, L. (2012). Optimal conservation policy under imperfect intergenerational altruism. Journal of Forest Economics 18, 194–206.

Di Maria, C., T. O. Michielsen, and E. van der Werf (2013). Carbon leakage. In J. F. Shogren (Ed.), Encyclopedia of Energy, Natural Resource, and Environmental Economics, Volume 2, pp. 255–259. Amsterdam: Elsevier.

Di Maria, C., S. Smulders, and E. van der Werf (2008). Absolute abundance and relative scarcity: Announced policy, resource extraction, and carbon emissions. FEEM Nota di Lavoro 92.2008.

Di Maria, C. and E. van der Werf (2012). Imperfect environmental policy and polluting emissions: The green paradox and beyond. International Review of Environmental and Resource Economics 6(2), 153–194.

Ehrlich, P. R. and A. Ehrlich (1981). Extinction. The causes and consequences of the disappearance of species. New York: Random House.

EIA (2000). The changing structure of the electric power industry 2000: An update. Technical report.

Eichner, T. and R. Pethig (2011). Carbon leakage, the green paradox and perfect future markets. International Economic Review 52(3), 767–805.

Ellingsen, T., M. Johannesson, J. Mollerstrom, and S. Munkhammar (2012). Social framing effects: Preferences or beliefs? Games and Economic Behavior 76(1), 117–130.

Ellison, G. and E. L. Glaeser (1999). The geographic concentration of industry: Does natural advantage explain agglomeration? American Economic Review: Papers and Proceedings 89(2), 311–316.

Felder, S. and T. F. Rutherford (1993). Unilateral CO2 reductions and carbon leakage: The consequences of international trade in oil and basic materials. Journal of Environmental Economics and Management 25(2), 162–176.

Fischer, C. and A. K. Fox (2009). Comparing policies to combat emissions leakage: Border tax adjustments versus rebates. RFF Discussion Paper 09-02-REV.

Fischer, C. and S. Salant (2010). On Hotelling, emissions leakage, and climate policy alternatives. Resources for the Future, Discussion Paper.

Fuji, T. and L. Karp (2006). Numerical analysis of non-constant discounting with an application to renewable resource management. CUDARE Working Paper 1019, Department of Agricultural and Resource Economics, UC Berkeley.

Gallini, N., T. Lewis, and R. Ware (1983). Strategic timing and pricing of a substitute in a cartelized resource market. Canadian Journal of Economics 16(3), 429–446.

Gattig, A. and L. Hendrickx (2007). Judgemental discounting and environmental risk perception: Dimensional similarities, domain differences, and implications for sustainability. Journal of Social Issues 63(1), 21–39.

Gerking, S. and S. F. Hamilton (2008). What explains the increased utilization of Powder River Basin coal in electric power generation? American Journal of Agricultural Economics 90(4), 933–950.

Gerlagh, R. (2011). Too much oil? CESifo Economic Studies 57(1), 79–102.

Gerlagh, R. and M. Liski (2011). Strategic resource dependence. Journal of Economic Theory 146(2), 699–727.

Gerlagh, R. and M. Liski (2012). Carbon prices for the next thousand years. CESifo Working Paper Series No. 3855.

Gerlagh, R. and N. A. Mathys (2011). Energy abundance, trade, and industry location. FEEM Nota di Lavoro 03.2011.

Gerlagh, R. and T. O. Michielsen (2013). The green paradox. In J. F. Shogren (Ed.), Encyclopedia of Energy, Natural Resource, and Environmental Economics, Volume 2, pp. 309–313. Amsterdam: Elsevier.

Golombek, R., M. Greaker, and M. Hoel (2010). Climate policy without commitment. CESifo Working Paper Series No. 2909.

Grafton, R. Q., T. Kompas, and N. V. Long (2012). Substitution between biofuels and fossil fuels: Is there a green paradox? Journal of Environmental Economics and Management 64(3), 328–341.

Greenley, D. A., R. G. Walsh, and R. A. Young (1981). Option value: Empirical evidence from a case study of recreation and water quality. Quarterly Journal of Economics 96(4), 657–673.

Grenadier, S. R. and N. Wang (2007). Investment under uncertainty and time-inconsistent preferences. Journal of Financial Economics 84(1), 2–39.

Hansen, G. D. and E. C. Prescott (2002). Malthus to solow. American Economic Review 92(4), 1205–1217.

Harberger, A. C. (1978). Perspectives on capital and technology in Less Developed Countries. In M. Artis and A. Nobay (Eds.), Contemporary Economic Analysis, pp. 15–40. London: Croom Helm.

Harris, C. and J. Vickers (1995). Innovation and natural resources: A dynamic game with uncertainty. RAND Journal of Economics 26(3), 418–430.

Harris, C. D. (1954). The market as a factor in the localization of industry in the United States. Annals of the Association of American Geographers 44(4), 315–348.

Harstad, B. (2012). Buy coal! a case for supply-side environmental policy. Journal of Political Economy 120(1), 77–115.

Helm, D., C. Hepburn, and R. Mash (2003). Credible carbon policy. Oxford Review of Economic Policy 19(3), 438–450.

Hepburn, C., S. Duncan, and A. Papachristodoulou (2010). Behavioural economics, hyperbolic discounting and environmental policy. Environmental and Resource Economics 46, 189–206.

Hillman, A. L. and C. W. Bullard (1978). Energy, the Heckscher-Ohlin theorem, and U.S. international trade. American Economic Review 68(1), 96–106.

Hlasny, V. (2006). Do gas cost incentives work? A nation-wide study. The American Economist 50(1), 51–68.

Hlasny, V. (2008). The impact of restructuring and deregulation on gas rates. Journal of Regulatory Economics 34(1), 27–52.

Hlasny, V. (2011). Economic determinants of deregulation in the gas distribution market. Journal of Economic Policy Reform 14(3), 201–213.

Hoel, M. (1978). Resource extraction, substitute production, and monopoly. Journal of Economic Theory 19(1), 28–37.

Hoel, M. (1983). Future conditions and present extraction: A useful method in natural resource economics. Resources and Energy 5(4), 303–311.

Hoel, M. (1994). Efficient climate policy in the presence of free riders. Journal of Environmental Economics and Management 27(3), 259–274.

Hoel, M. (2010). Is there a Green Paradox? CESifo Working Paper 3168.

Hoel, M. (2011). The supply side of CO2 with country heterogeneity. Scandinavian Journal of Economics 113(3), 846–865.

Hoel, M. and S. Kverndokk (1996). Depletion of fossil fuels and the impacts of global warming. Resource and Energy Economics 18(2), 115–136.

IEA (2010a). Coal information 2010.

IEA (2010b). World Energy Outlook.

IEA (2011). World Energy Outlook.

IPCC (2007). Contribution of working group i to the fourth assessment report of the intergovernmental panel on climate change.

Jaakkola, N. (2012). Strategic oil supply and gradual development of substitutes. Mimeographed.

Jevons, W. S. (1865). The Coal Question; An Inquiry Concerning the Progress of the Nation, and the Probable Exhaustion of Our Coal Mines. London: Macmillan & Co.

Joskow, P. L. (2000). Deregulation and regulatory reform in the U.S. electric power sector. In S. Peltzman and C. Winston (Eds.), Deregulation of Network Industries: What's Next?, pp. 113–188. Washington, DC: Brookings Institution Press.

Kaoru, Y. (1993). Differentiating use and nonuse values for coastal pond water quality improvements. Environmental and Resource Economics 3, 487–494.

Karp, L. (2005). Hyperbolic discounting and climate change. Journal of Public Economics 89(2-3), 261–282.

Karp, L. and D. M. Newbery (1993). Intertemporal consistency issues in depletable resources. In A. V. Kneese and J. L. Sweeny (Eds.), Handbook of Natural Resource and Energy Economics, Volume 3. Elsevier.

Karp, L. and Y. Tsur (2011). Time perspective and climate policy. Journal of Environmental Economics and Management 62(1), 1–14.

Keller, K., B. M. Bolker, and D. F. Bradford (2004). Uncertain climate thresholds and optimal economic growth. Journal of Environmental Economics and Management 48(1), 723–741.

Kharecha, P. A. and J. E. Hansen (2008). Implications of "peak oil" for atmospheric CO2 and climate. Global Biogeochemical Cycles 22(3), GB3012.

Ko, J. and C. Dahl (2001). Interfuel substitution in US electricity generation. Applied Economics 33(14), 1833–1843.

Krusell, P., B. Kuruşçu, and A. Smith (2002). Equilibrium welfare and government policy with quasi-geometric discounting. Journal of Economic Theory 105, 42–72.

Krusell, P. and A. Smith (2003). Consumption-savings decisions with quasi-geometric discounting. Econometrica 71(1), 365–375.

Kuik, O. and M. Hofkes (2010). Border adjustment for European emissions trading: Competitiveness and carbon leakage. Energy Policy 38(4), 1741–1748.

Laibson, D. (1998). Life-cycle consumption and hyperbolic discount functions. European Economic Review 42(3-5), 861–871.

Langford, I. H., A. Kontogianni, M. S. Skourtos, S. Georgiou, and I. J. Bateman (1998). Multivariate mixed models for open-ended contingent valuation data: Willingness to pay for conservation of monk seals. Environmental and Resource Economics 12, 443–456.

Layton, D. and R. Levine (2003). How much does the far future matter? a hierarchical bayesian analysis of the public's willingness to mitigate ecological impacts of climate change. Journal of the American Statistical Association 98(463), 533–544.

Layton, D. F. and G. Brown (2000). Heterogeneous preferences regarding global climate change. The Review of Economics and Statistics 82(4), 616–624.

Lenton, T. M., H. Held, E. Kriegler, J. W. Hall, W. Lucht, S. Rahmstorf, and H. J. Schellnhuber (2008). Tipping elements in the Earth's climate system. Proceedings of the National Academy of Sciences of the United States of America 105(6), 1786–1793.

Li, C.-Z. and K.-G. Löfgren (2000). Renewable resources and economic sustainability: A dynamic analysis with heterogeneous time preferences. Journal of Environmental Economics and Management 40(3), 236–250.

Liski, M. and J.-P. Montero (2011). On the exhaustible-resource monopsony. Mimeographed.

Long, N. and V. Martinet (2012). Combining rights and welfarism: A new approach to intertemporal evaluation of social alternatives. CESifo Working Paper Series No. 3746.

Ma, H., L. Oxley, J. Gibson, and B. Kim (2008). China's energy economy: Technical change, factor demand and interfactor/interfuel substitution. Energy Economics 30(5), 2167–2183.

Michaels, G. (2010). The long term consequences of resource-based specialisation. The Economic Journal 121(551), 31–57.

Michielsen, T. O. (2011). Brown backstops versus the green paradox. CentER Discussion Paper Series No. 2011-110.

Michielsen, T. O. (2013a). The distribution of energy-intensive sectors in the USA. Journal of Economic Geography. http://dx.doi.org/10.1093/jeg/lbs045.

Michielsen, T. O. (2013b). Environmental catastrophes under time-inconsistent preferences. CentER Discussion Paper Series No. 2013-013.

Michielsen, T. O. (2013c). Strategic resource extraction and substitute development. CentER Discussion Paper Series No. 2013-014.

Midelfart-Knarvik, K. H., H. G. Overman, and A. J. Venables (2000). Comparative advantage and economic geography: Estimating the location of production in the EU. CEPR discussion paper no. 2618.

Millner, A. (2013). On welfare frameworks and catastrophic climate risks. Journal of Environmental Economics and Management. http://dx.doi.org/10.1016/j.jeem.2012.09.006.

Miranda, M. J. and P. L. Fackler (2002). Applied Computational Economics and Finance. Cambridge, MA: MIT Press.

Mulatu, A., D. Rigby, R. Gerlagh, and A. Wossink (2010). Environmental regulation and industry location in Europe. Environmental and Resource Economics 45, 459–479.

Muradian, R. (2001). Ecological thresholds: a survey. Ecological Economics 38, 7–24.

Nævdal, E. (2006). Dynamic optimisation in the presence of threshold effects when the location of the threshold is uncertain - with an application to the possible disintegration of the Westerm Antarctic Ice Sheet. Journal of Economic Dynamics and Control 30(7), 1131–1158.

Nepstad, D. C., C. M. Stickler, B. Soares-Filho, and F. Merry (2008). Interactions among Amazon land use, forests and climate: prospects for a near-term forest tipping point. Philosophical Transactions of the Royal Society B 363(1498), 1737–1746.

Nordhaus, W. D. (1994). Managing the Global Commons: The Economics of Climate Change. Cambridge, MA: MIT Press.

Nunes, P. A. L. D. and J. C. J. M. van den Bergh (2001). Economic valuation of biodiversity: sense or nonsense? Ecological Economics 39, 203–222.

O'Donoghue, T. and M. Rabin (1999). Doing it now or later. American Economic Review 89(1), 103–124.

Olsen, T. E. (1986). Strategic considerations in invention and innovation: The case of natural resources revisited. Econometrica 56(4), 841–849.

Olsen, T. E. (1993). Perfect equilibrium timing of a backstop technology. Journal of Economic Dynamics and Control 17(1-2), 123–151.

Paltsev, S. V. (2001). The Kyoto Protocol: regional and sectoral contributions to the carbon leakage. The Energy Journal 22(4), 53–79.

Papyrakis, E. and R. Gerlagh (2007). Resource abundance and economic growth in the United States. European Economic Review 51(4), 1011–1039.

Perkins, F. C. (1994). A dynamic analysis of Japanese energy policies: their impact on fuel switching and conservation. Energy Policy 22(7), 595–607.

Persson, T. A., C. Azar, D. Johansson, and K. Lindgren (2007). Major oil exporters may profit rather than lose, in a carbon-constrained world. Energy Policy 35(12), 6346–6353.

Polasky, S., A. de Zeeuw, and F. Wagener (2011). Optimal management with potential regime shifts. Journal of Environmental Economics and Management 62(2), 229–240.

Ray, G. F. (1979). Energy economics - a random walk in history. Energy Economics 1(3), 139–143.

Reed, W. J. (1984). The effects of the risk of fire on the optimal rotation of a forest. Journal of Environmental Economics and Management 11(2), 180–190.

Romalis, J. (2004). Factor proportions and the structure of commodity trade. American Economic Review 94(1), 67–97.

Scheffer, M. (1997). Ecology of Shallow Lakes. London: Chapman and Hall.

Schott, P. K. (2003). One size fits all? Heckscher-Ohlin specialization in global production. American Economic Review 93(3), 686–708.

Serletis, A., G. Timilsina, and O. Vasetsky (2009). On interfuel substitution: Some international evidence. World Bank Policy Research Working Paper 5026.

Serletis, A., G. R. Timilsina, and O. Vasetsky (2010). Interfuel substitution in the United States. Energy Economics 32(3), 737–745.

Simon, L. K. and M. B. Stinchcombe (1989). Extensive form games in continuous time: Pure strategies. Econometrica 57(5), 1171–1214.

Sims, R. E. H., W. Mabee, J. N. Saddler, and M. Taylor (2010). An overview of second generation biofuel technologies. Bioresource Technology 101(6), 1570–1580.

Sinn, H.-W. (2008a). Das grüne Paradoxon, Plädoyer für eine illusionsfreie Klimapolitik. Econ-Verlag.

Sinn, H.-W. (2008b). Public policies against global warming: a supply side approach. International Tax and Public Finance 15(4), 360–394.

Smulders, S. and E. van der Werf (2008). Climate policy and the optimal extraction of high- and low-carbon fossil fuels. Canadian Journal of Economics 41(4), 1421–1444.

Söderholm, P. (2000). Environmental regulations and interfuel substitution in the power sector: A generalized Leontief model. Energy & Environment 11(1), 1–23.

Söderholm, P. (2001). Fossil fuel flexibility in west European power generation and the impact of system load factors. Energy Economics 23(1), 77–97.

Stern, D. I. (2012). Interfuel substitution: A meta-analysis. Journal of Economic Surveys 26(2), 307–331.

Stern, N. (2007). The Economics of Climate Change: The Stern Review. Berlin: Cambridge University Press.

Strand, J. (2007). Technology treaties and fossil-fuels extraction. The Energy Journal 28(4), 129–142.

Stutz, J. (2010). The three-front war: Pursuing sustainability in a world shaped by explosive growth. Sustainability: Science, Practice, & Policy 6(2), 49–59.

Trefler, D. (1995). The case of missing trade and other mysteries. American Economic Review 85(5), 1029–1046.

Tsur, Y. and A. Zemel (1994). Endangered species and natural resource exploitation: Extinction vs. coexistence. Natural Resource Modeling 8(4), 389–413.

Tsur, Y. and A. Zemel (1996). Accounting for global warming risks: Resource management under event uncertainty. Journal of Economic Dynamics and Control 20(6), 1289–1305.

Tsur, Y. and A. Zemel (2008). Regulating environmental threats. Environmental and Resource Economics 39, 297–310.

van der Ploeg, F. and C. Withagen (2012a). Is there really a green paradox? Journal of Environmental Economics and Management 64(3), 342–363.

van der Ploeg, F. and C. Withagen (2012b). Too much coal, too little oil. Journal of Public Economics 96(1-2), 62–77.

Warde, P. (2006). Fear of wood shortage and the reality of the woodland in Europe, c. 1450-1850. History Workshop Journal 62(1), 28–57.

Wattage, P. and S. Mardle (2008). Total economic value of wetland conservation in Sri Lanka identifying use and non-use values. Wetlands Ecology and Management 16(5), 359–369.

Weitzman, M. (2001). Gamma discounting. American Economic Review 91(1), 260–271.

Weitzman, M. (2009). On modeling and interpreting the economics of catastrophic climate change. Review of Economics and Statistics 91(1), 1–19.

Weitzman, M. (2010). GHG targets as insurance against catastrophic climate damages. Mimeographed.